

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2003, 14:37:09 ; Search time 38 Seconds
(without alignments)
4948.535 Million cell updates/sec

Title: US-10-021-571-4

Perfect score: 8884
Sequence: 1 MSVLISOSVINYVEENIPN.....LHAASSESTGGERESTIL 1715

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 424699 seqs, 10964683 residues

Total number of hits satisfying chosen parameters: 424699

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCRT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3050.5	34.3	705	US-10-106-698-6378	Sequence 6378, Ap
2	2851.5	32.1	551	US-09-835-788A-17	Sequence 17, Appl
3	2638	29.7	513	US-10-149-819-9	Sequence 9, Appl
4	553	6.2	109	US-09-986-480-371	Sequence 371, App
5	519	5.8	1724	US-09-964-899-43	Sequence 43, Appl
6	437	4.9	740	US-09-835-788A-12	Sequence 12, Appl
7	408.5	4.6	426	US-09-908-711-70	Sequence 70, Appl
8	376.5	4.2	285	US-09-835-788A-18	Sequence 18, Appl
9	376.5	4.2	1327	US-09-972-115A-8	Sequence 8, Appl
10	376.5	4.2	1327	US-09-841-835-2	Sequence 2, Appl
11	370.5	4.2	949	US-09-841-835-10	Sequence 10, Appl
12	370	4.2	1166	US-09-972-115A-6	Sequence 6, Appl
13	370	4.2	1166	US-10-163-587A-15	Sequence 15, Appl
14	357	4.0	1074	US-09-163-587A-2	Sequence 2, Appl
15	348.5	3.9	784	US-10-164-080-7	Sequence 7, Appl
16	347	3.9	802	US-09-964-899-41	Sequence 41, Appl
17	346	3.9	251	US-09-835-788A-13	Sequence 13, Appl
18	345.5	3.9	1333	US-09-972-115A-2	Sequence 2, Appl
19	344	3.9	786	US-10-164-080-2	Sequence 2, Appl

20	344	3.9	786	US-10-299-327-2	Sequence 2, Appl
21	344	3.9	787	US-09-866-050A-334	Sequence 334, App
22	341.5	3.8	679	US-10-339-936-2	Sequence 2, Appl
23	339.5	3.8	1267	US-09-972-115A-4	Sequence 4, Appl
24	320	3.6	835	US-09-947-199-2	Sequence 2, Appl
25	309.5	3.5	673	US-09-841-835-8	Sequence 8, Appl
26	293	3.3	835	US-09-841-835-8	Sequence 8, Appl
27	288.5	3.2	306	US-09-835-788A-19	Sequence 19, Appl
28	285	3.2	1719	US-10-012-896-378	Sequence 378, App
29	285	3.2	1719	US-09-895-793-378	Sequence 378, App
30	285	3.2	1719	US-09-895-814-378	Sequence 378, App
31	285	3.2	1719	US-10-010-940-378	Sequence 378, App
32	285	3.2	1719	US-09-759-143-378	Sequence 378, App
33	285	3.2	1719	US-09-780-669-378	Sequence 378, App
34	285	3.2	1719	US-09-822-827-378	Sequence 378, App
35	265.5	3.0	599	US-09-735-368-2	Sequence 2, Appl
36	265.5	2.9	622	US-10-197-666A-22	Sequence 22, Appl
37	238.5	2.7	622	US-10-197-666A-6	Sequence 6, Appl
38	235	2.6	329	US-09-880-192-62	Sequence 62, Appl
39	235	2.6	329	US-09-758-593A-1	Sequence 1, Appl
40	235	2.6	1054	US-09-798-042-87	Sequence 87, Appl
41	234	2.6	1349	US-10-076-622-573	Sequence 573, App
42	234	2.6	1349	US-10-007-805-573	Sequence 573, App
43	232	2.6	399	US-10-076-622-572	Sequence 572, App
44	232	2.6	399	US-10-007-805-572	Sequence 572, App
45	231.5	2.6	328	US-09-758-593A-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1
US-10-106-698-6378
Sequence 6378, Application US/10106698
Publication No. US20030109690A1
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypept
FILE REFERENCE: PA005P1
CURRENT FILING DATE: 2002-03-27
CURRENT FILING DATE: 2002-03-27
PRIOR APPLICATION NUMBER: PCT/US00/26524
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US 60/157,137
PRIOR FILING DATE: 1999-09-29
PRIOR APPLICATION NUMBER: US 60/163,280
PRIOR FILING DATE: 1999-11-03
NUMBER OF SEQ ID NOS: 8564
SOFTWARE: PatentIn Ver. 3.0
SEQ ID NO 6378
LENGTH: 705
TYPE: PR
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (244)
NAME/KEY: MISC FEATURE
LOCATION: (337)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-6378
Query Match 34.3%; Score 3050.5; DB 9; Length 705;
Best Local Similarity 98.8%; Pred. No. 3.1e-176;
Matches 587; Conservative 2; Mismatches 4; Indels 1; Gaps 1;
DB 88 LILSHGANPSVIGL YSYVPIIIMAGRGHADIIVHLLDNGAKVNCSDKYGTPPLVMAARK 146
122 LILSHGANPSVIGL YSYVPIIIMAGRGHADIIVHLLDNGAKVNCSDKYGTPPLVMAARK 181
62 ANCNEDLDNMTALISASKEGHVHYVEELKGVNLEHDMGWTALMACKGRDVE 121
28 SNCNEDLDNMTALISASKEGHVHYVEELKGVNLEHDMGWTALMACKGRDVE 87

182 GHECYKHLAMGADVDOEGANSMTALIVAVGCTTQSVKEILKRNPNVNLTDKGNAL 241
147 GHECYKHLAMGADVDOEGANSMTALIVAVGCTTQSVKEILKRNPNVNLTDKGNAL 206
242 MIASKRGHEIYODLLDAGTYVNIIPDRSGDTVLIGAVRGHEIYRALLQKADIDINGO 301
207 MIASKRGHEIYODLLDAGTYVNIIPDRSGDTVLIGAVRGHEIYRALLQKADIDINGO 266
302 DKKTALYAVVEKGNATWVDILQCNPDTEICTKDETPLIKATKRNIEVEVELLDKAK 361
267 DKKTALYAVVEKGNATWVDILQCNPDTEICTKDETPLIKATKRNIEVEVELLDKAK 326
362 VSAVKKGGTPIHAIIRGSRKLAELLRNPKDGLLRPNKAGTTPNIDCSHOKSLT 421
327 VSAVKKGGTPIHAIIRGSRKLAELLRNPKDGLLRPNKAGTTPNIDCSHOKSLT 386
422 QIFGARHLSPETDGMGLGYDLYSSALADILSEPTMOPPICVGLYQWGSKSFLLKLE 481
387 QIFGARHLSPETDGMGLGYDLYSSALADILSEPTMOPPICVGLYQWGSKSFLLKLE 446
482 DEMKTFAGQOIEPLFQFMSLVFLLTLLCGGLLFAFTVHPNIGIANSLSFLALLYTF 541
447 DEMKTFAGQOIEPLFQFMSLVFLLTLLCGGLLFAFTVHPNIGIANSLSFLALLYTF 506
542 IVIYGGRRGESSNMWAVLSTRLARHIGYLELLKLMFVNPPELPEOTTKALPVRLFT 601
507 IVIYGGRRGESSNMWAVLSTRLARHIGYLELLKLMFVNPPELPEOTTKALPVRLFT 566
602 DYNRLSSVGETSLAEMATLSDACEREFGLARLRFVREKTEDTQKKKKKT 655
567 DYNRLSSVGETSLAEMATLSDACEREFGLARLRFVREKTEDTQKKKKKT 620

RESULT 2

US-09-835-788A-17
; Sequence 17, Application US/09835788A
; Patent No. US20020077458A1
; GENERAL INFORMATION:
; APPLICANT: NI et al.
; TITLE OF INVENTION: Death Domain-Containing Receptor Polynucleotides, Polypeptides,
; FILE OF INVENTION: Antibodies
; FILE REFERENCE: PTO18P1
; CURRENT APPLICATION NUMBER: US/09/835,788A
; PRIOR FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: PCT/US00/28666
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/159,585
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 60/167,246
; PRIOR FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 17
; LENGTH: 551
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-835-788A-17

Query Match 32.1% Score 2851.5; DB 10; Length 551;
Best Local Similarity 99.6%; Pred. No. 2.3e-164;
Matches 550; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

102 MGGTALMAYCYGRDYNVLLSHGANSVYGLQTSVPTIWAAGRGADIVHLLONG 161
1 MGGTALMAYCYGRDYNVLLSHGANSVYGLQTSVPTIWAAGRGADIVHLLONG 59
162 AKVNCSDKYGTPLVMAARKGHECYKHLAMGADVDOEGANSMTALIVAVGCTTQSV 221
60 AKVNCSDKYGTPLVMAARKGHECYKHLAMGADVDOEGANSMTALIVAVGCTTQSV 119
222 EILKRNPNVNLTDKGNATLMIASKRGHEIYODLLDAGTYVNIIPDRSGDTVLIGAVRG 281

120 EILKRNPNVNLTDKGNATLMIASKRGHEIYODLLDAGTYVNIIPDRSGDTVLIGAVRG 179
282 HVEIYRALLQKADIDIRGDNKPTALYAVVEKGNATWVDILQCNPDTEICTKDETPLI 341
180 HVEIYRALLQKADIDIRGDNKPTALYAVVEKGNATWVDILQCNPDTEICTKDETPLI 239
342 KATKRNIEVEVELLDKAKAVSAVKKGGTPIHAIIRGSRKLAELLRNPKDGLLRPN 401
240 KATKRNIEVEVELLDKAKAVSAVKKGGTPIHAIIRGSRKLAELLRNPKDGLLRPN 299
402 NKAGETTPNIDCSHOKSILQIFGARHLSPETDGMGLGYDLYSSALADILSEPTMOP 359
300 NKAGETTPNIDCSHOKSILQIFGARHLSPETDGMGLGYDLYSSALADILSEPTMOP 326
462 CVGLYQWGSKSFLLKLEDEMKTFAQOIEPLFQFMSLVFLLTLLCGGLLFAFTV 521
360 CVGLYQWGSKSFLLKLEDEMKTFAQOIEPLFQFMSLVFLLTLLCGGLLFAFTV 419
522 HPNIGIANSLSFLALLYTFIYVYGGRRGESSNMWAVLSTRLARHIGYLELLKLMF 581
420 HPNIGIANSLSFLALLYTFIYVYGGRRGESSNMWAVLSTRLARHIGYLELLKLMF 479
582 NPPELPEOTTKALPVRLFTDYNRLSSVGETSLAEMATLSDACEREFGLARLRFV 641
480 NPPELPEOTTKALPVRLFTDYNRLSSVGETSLAEMATLSDACEREFGLARLRFV 539
642 KTEDTQKKKKKT 653
540 KTEDTQKKKKKT 551

RESULT 3

US-10-149-819-9
; Sequence 9, Application US/10149819
; Publication No. US20030044913A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: YVE, Henry
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: TANG, Y. Tom
; APPLICANT: PATTERSON, Chandra
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: LU, Dying Anna M.
; APPLICANT: SHAH, Purvi
; APPLICANT: LAL, Preeti
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: BUREFORD, Neil
; TITLE OF INVENTION: EXTRACELLULAR MATRIX AND CELL ADHESION MOLECULES
; FILE REFERENCE: PF-0760 PCT
; CURRENT APPLICATION NUMBER: US/10/149,819
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/172,852; 60/172,354
; PRIOR FILING DATE: 1998-12-10; 1999-12-16
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL Program
; SEQ ID NO 9
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20030044913A1 2215706CD1
US-10-149-819-9

Query Match 29.7% Score 2638; DB 9; Length 513;
Best Local Similarity 99.8%; Pred. No. 1.7e-151;
Matches 512; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1203 MNMFGDWHLFRSTYLENRNASHVYVPEPFLSESSGPAHPGEPARRASHNELPTEL 1262
1 MNMFGDWHLFRSTYLENRNASHVYVPEPFLSESSGPAHPGEPARRASHNELPTEL 60
1263 SSQPTTILNFSPEELNIGLDEGAPRHSNLSWSQOTRTPSLSSLSNDSISLKLTK 1322

```

Db      61 SSGPTPTLNPFELNLTGIDEBAPRHSNLSMWSQTRRPTSLSLNSQSSSTIEISLTIDK 120
QY      1323 VQAEYDAYREYIAQMSQLEGFGSTTIGRSSPHSTYVMGSSGSGSIHNSLEDEKGD 1382
Db      121 VQAEYDAYREYIAQMSQLEGFGSTTIGRSSPHSTYVMGSSGSGSIHNSLEDEKGD 180
QY      1383 SEPKPDGKRSFLMKRGDVIYSSSGVSTNDASPLDPIEEDERKSQSSSKILLPGKSSSE 1442
Db      181 SEPKPDGKRSFLMKRGDVIYSSSGVSTNDASPLDPIEEDERKSQSSSKILLPGKSSSE 240
QY      1443 RSSLPQTDLLKSGSLRYOKLPDSEDESGTEESDNTPLIKDDKRAEKKVEYRPSPH 1502
Db      241 RSSLPQTDLLKSGSLRYOKLPDSEDESGTEESDNTPLIKDDKRAEKKVEYRPSPH 300
QY      1503 SAEPIRTFKAKYELSDALLDKKSSDGVRSSESSPNHSLNEVADDSHLEKANLIELE 1562
Db      301 SAEPIRTFKAKYELSDALLDKKSSDGVRSSESSPNHSLNEVADDSHLEKANLIELE 360
QY      1563 DSHSGRGIPHSLSGLDPIIARMGICSEDKKSPSECSLIASSPENMPACOKAYNLNR 1622
Db      361 DSHSGRGIPHSLSGLDPIIARMGICSEDKKSPSECSLIASSPENMPACOKAYNLNR 420
QY      1623 TPSTVTLNNSAPANRANONFDEMEGIRTSQYILRPSSSPNPTTIGNENLKSMTKRSG 1682
Db      421 TPSTVTLNNSAPANRANONFDEMEGIRTSQYILRPSSSPNPTTIGNENLKSMTKRSG 480
QY      1683 RSSYTRLSKDPPELHAAASSTGFGBERESIL 1715
Db      481 RSSYTRLSKDPPELHAAASSTGFGBERESIL 513

```

RESULT 4

```

US-09-986-480-371
; Sequence 371, Application US/09986480
; Publication No. US20030027999A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 143 Human Secreted Proteins
; FILE REFERENCE: P8500P1
; CURRENT APPLICATION NUMBER: US/09/986,480
; CURRENT FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: PCT/US00/12788
; PRIOR FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: US 60/134,068
; PRIOR FILING DATE: 1999-05-13
; NUMBER OF SEQ ID NOS: 456
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 371
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-986-480-371

```

```

Query Match      6.2%; Score 553; DB 9; Length 109;
Best Local Similarity 100.0%; Pred. No. 2e-26;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      302 DNKTLTYWAVEKGNAYNRDILQCNPDTEICTGDETPLIKATKRNIEVEILLDKGAK 361
Db      1 DNKTLTYWAVEKGNAYNRDILQCNPDTEICTGDETPLIKATKRNIEVEILLDKGAK 60
QY      362 VSAADKRGDTPHAIIRGRSRKLAELLRNPKDGRILYRPNKAGETP 408
Db      61 VSAADKRGDTPHAIIRGRSRKLAELLRNPKDGRILYRPNKAGETP 107

```

```

RESULT 5
US-09-964-899-43
; Sequence 43, Application US/09964899
; Patent No. US2002017446A1
; GENERAL INFORMATION:
; APPLICANT: Cohen, Dalila et al.

```

```

; TITLE OF INVENTION: Identification of Genes Involved in
; FILE REFERENCE: 4-31612 A
; CURRENT APPLICATION NUMBER: US/09/964,899
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/236,893
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/298,309
; PRIOR FILING DATE: 2001-06-14
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 1724
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-964-899-43

```

Query Match 5.8%; Score 519; DB 9; Length 1724;

Best Local Similarity 20.5%; Pred. No. 1.3e-22;

Matches 383; Conservative 251; Mismatches 663; Indels 568; Gaps 69;

```

QY      14 EENIPALKALLECKDVERNEGCOTPLMAIEGNGLEIYKELIKNGANCNLEDUNWT 73
Db      45 KEGHEVVESELLQREANVDAATKKGNTALHIASLAGAEVYKVTNGANVNAOSQNGFT 104
QY      74 ALISAKKEGVHIVEELKCGVNLBHRMGWTALMMCYGRDVEILL- 124
Db      105 PLYMAOENHLEVAFLDNGASQSLATEDGFTPLVALQOCHDOVVELLENDTKKVR 164
QY      125 -----SHGANSVYGL- 135
Db      165 LPAHLIAARKDPTKAALLLQNDNNADYESKSGFTPLHIAHYGINVATLLNRAAVD 224
QY      136 ---QYSVPTIWAAGRHADIVHLLONGAKVNCSDKGTTPLYAAKRG- 182
Db      225 PTAARDITPLHVAASKRGNAMVTKLLIDNRKAKIDAATRGITPLPHCGASGHEQVVEMLD 284
QY      183 -----HLECYVHLLAMGADVOEGANSMTALIYAVKGYTQS 219
Db      285 RAAPILSKTKNGLSPLHMAATOGDHLNCGVOLLQHNVPDDVDYNDVITPLHVAHAGHYKV 344
QY      220 VREILKRNPNVNLTRKDGNTALMAISKEGHEIYVDLLDAGTYVNIIPRSGDTVLIGAVR 279
Db      345 AVILDKKANPNKAKNLGFTPLHICKNRKRYMELLKHGASIQAVTERGETALHMAAR 404
QY      280 GGHVEIVRALLOKYADIDIRGODNKTALYMAVEKGNAYNRDILQCNPDTEICTGDETP 339
Db      405 SGQAEVRYLYVDGAQVEKAKDDQTPHISARLGRADIVQOLLQOGASPNAAATTSGYTP 464
QY      340 LKATKRNIEVEILLDKGAKVSAVDKGDPTPLHAIIRGRSRKLAELLRNPKDGRILY 399
Db      465 LHSAREGHEVDAAFLDHDGASLSTTKSGITPLHVAHAYDNOKALLL- DQGASPH 521
QY      400 RPNKAGETPNIDC-SHOKSILTOI--FGAR-----HLSPTETDGMGLGYLY 444
Db      522 AAANKGYTPLHIAAKKNQNDIATLLEYGADANAATYRGQIASVHLAAEGHNDVSLLDG 581
QY      445 SSALADILSEPTMOPICVGLYAQMSGSGSFLKLIED-----EKTATAGQOIE- 493
Db      582 RNNANVNLMSKSGTLP--LHLAAQ-----EDVNVVAEVLVNOGAVHDAQOTYV 626
QY      494 --PLFOFSLVIELFTLLCG-----LGLFAFYTPNGLIVASLFTAL 536
Db      627 GPPLPGRKCVHLYT---ANGTTPHQAAQOQHTHINVLQNNASPN-ELVYVTEKHK 682
QY      537 LYIEFTVIVYGGREGESNMAMVLSRILARIIGYELILKIMFVN-----P 584
Db      683 MNV-----PETMNEVYLDMSDDDEVKRNANPEMLSDXEYISDVEEGNRCWTYKIP 730
QY      585 ELPEQTTKALPVAFLETDVNNRLSSVGGESLSLPMATLSDAEREGFLATLFFVFKTE 644
Db      731 KVOEFTVKDTPFRREAFDVLSTLSAGEBAM-----TG 763

```

QY 645 DTQ--GKKMK--TCCLPSEVIFLEIIGCIISGTTLLAIFRVDKHLTVNAVLISAS 699
 764 DFDKYLGPDLKELGDDSLPABGYMGFSIGA--RSARLPSFMVABGSMNG-----SR 816
 QY 700 VGLAVL--NCRWMOYVDSLNSQKRLN-----AASKLKLKSEG----- 741
 817 HHGMRLLIPRCKTAPRITCRLY--KRHLKLANPPMVBEGIASLVEIMGAGAOFLCP 874
 QY 742 -----FMKVLKCEVLMARAKTIDSTFONO-----TRLVIIIDJLADAGQDVLQML 789
 875 VIVEIPFGSMGKEHLLVLSNGETKEHOFOSKNDLLELLGMD-----EEL 926
 QY 790 DTVRVLK--GPIALFASD----- 808
 927 DSPEELGKKRICRIITKDPQYFAVYSRIKESNOGPGGILSTTVLVQASPEGAL 986
 QY 809 -----PHIILKAINONLSVARDNINHDMYIVLPL-----NSR 848
 987 TKRIRVGLAOVPDEIVAKI--LGNKATFSPITVEFRKRRKFKPITMTIPVPPSGE 1043
 QY 849 GLSNARK-----FLVTSATNGDVP--SDTTGIOEDADRVRSONSLCEMTKLGSKTAL 899
 1044 GYSNGYKGDPTNLRLCISITGTSAPQMEDITG--TTPLFIDCVSFTTNVSR--- 1097
 QY 900 NRDYTRRROMQRTTROMSFDLTKLVTEDWFSDISPTQMRLLNIVSVTGRLLRANQI 959
 1098 -----YGNKGPOKAV-----LEKPIYVDCYGNLAPLTK-----GGOQLVFNFY 1136
 QY 960 SFMDRLASMINL--TEOWPYRTSMJLLEETEGIPD----- 995
 1137 SKRENLPISIKIRTSOEPGRSLFKPKTKTKGIPOTAVCNLITLPAHKIKETDRR 1196
 QY 996 ---QMTLTYERISKNIPTTKDVEPL--EIDGDIRNEVELSR-----TPVLVARD 1044
 1197 OSFASIALKRRYSYLTEPGMKKKWQSELSDEESTSNTLSLERSRGGQPSVTTKSARD 1256
 QY 1045 VAVFLPCTYNLDPKLEITADVRAREQI-----STGG 1077
 1257 KYTE--AAPLKSEKSEKSSRRADALSVLTKINRIDITYLEGPIFDYGNISG 1313
 QY 1078 L-----AYP-----PLPHEGPPRPSGYSOPSCSSFMGPPAG 1114
 1314 TSPADENNVEHDPVGYSLQVELETPTGLHYTP--PTPQODDYSDISLISPL-- 1368
 QY 1115 GYVSPQPHSSYSYSGMTGQHPFYNRGSGPAPGVVLLNSLVANDAVCEKLIKQIEGLDQSM 1174
 1369 -----RTPSRLSDGLVPSQGNIEHSADGP--PVYTAE--DASLEDSKLEDSVPLEM 1416
 QY 1175 PQYCTTIRKANINGVLAQCNIDELKEMNMFGWMLFRSTVLEMRNAESHVVPEDRF 1234
 1417 PE-----AVDDESQLE--NVCLSSEYPOYL--NLASGPPDVKPAEPRK 1456
 QY 1235 LSESSGAPHGEPARRASHNELPTELSQTPYTLNFSFELNLTGLDEGAPRRHSNLSW 1294
 1457 LGVSE--QOEKKSQGPDEEMMEKIKS-----LFEDIQ--LEEG-----YES 1495
 QY 1295 OSQTRTSPSLNSQSSIEISKLTDVQVAREYRAYEYIAQMSQLEGGPSTTISR- 1353
 1496 EEMTEEKYQALIKRVOQAELEMSITG--WONETSSGNLESCAQARVYRG-----LIDRL 1549
 QY 1354 -SSPH-----STYMGSSSGSGSIHSLNLEQEKGDSPKDDGKSKSLM--RGDVIDYS 1405
 1550 DDPQOCHDSTITSYKGEA--GKEAN-----GSHTEITPEAKTKSYFPESONDVYKOS 1601
 QY 1406 S-----SGVSTNASPPLDPTTEDEKSDQSGSKL-----PGKSSERSSLF 1447
 1602 TKETLKPKIHSGHGEERASPLAAV-----QKSLETSKLLIEBKPCVPDLKDESDSSS 1657
 QY 1448 QTDLKSGSLKYOKLPSDED-----ESGTEESDNTPLLDK--DRKAEKVERVP 1497
 1658 EEBRRTVTRVIRRRLLIKGEBAKNIPGSEVTEQ-----FTDEGNLITRKGEQFVKYTK 1712
 QY 1498 KSPER 1502

Db 1713 KEIRH 1717
 RESULT 6
 US-09-835-788A-12
 ; Sequence 12, Application US/09835788A
 ; Patent No. US20020077458A1
 GENERAL INFORMATION:
 ; APPLICANT: NI et al.
 ; TITLE OF INVENTION: Death Domain-Containing Receptor Polynucleotides, Polypeptides
 ; FILE REFERENCE: PT018P1
 ; CURRENT FILING DATE: 2001-04-17
 ; PRIOR APPLICATION NUMBER: US/09/835,788A
 ; PRIOR FILING DATE: 2000-10-17
 ; PRIOR APPLICATION NUMBER: PCT/US00/28666
 ; PRIOR FILING DATE: 2000-10-17
 ; PRIOR APPLICATION NUMBER: 60/159,585
 ; PRIOR FILING DATE: 1999-10-18
 ; PRIOR APPLICATION NUMBER: 60/167,246
 ; PRIOR FILING DATE: 1999-11-24
 ; NUMBER OF SEQ ID NOS: 24
 ; SOFTWARE: Patent Ver. 2.0
 ; SEQ ID NO 12
 ; LENGTH: 740
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-835-788A-12

Query Match 4.98; Score 437; DB 10; Length 740;

Best Local Similarity 30.3%; Pred. No. 3,4e-18;

Matches 147; Conservative 66; Mismatches 192; Indels 80; Gaps 13;

QY 6 SOSVYNVEENIPALKALEKCKDVERREC-QPPLMAAOGNLEIKELKNGANC 64
 303 SQTSNVASQSMRPYPSV-----DIDAHESNHDALTLACAGHEEIVSLIARDAKI 357
 QY 65 NLEDLMDTALLISAKEGHVIIEELKCGVNE--HRDMGWTALMAGYKRTDVELL 123
 358 EHRDKKGFPLLAATAGHVGVEILLDKGDIENQSERKQDPLSLACSGGQGEVVDLL 417
 QY 124 LSHGANPSVGLQTSYPTTAAAGRHADIIVHLLONGAKVN--CSDKGTGTPLVAAK 181
 418 LARGANEHRNVS--DYPLSLAASGGVYNLIKILNAGAEINRSRTSGKISPLMAAMN 476
 QY 182 GHLECYKHLAMGADV--OEGANSMTALIVAVGGTOSYKELTKGNPNVNLTDKXGNTA 240
 477 GHYPAVAYLLDMGSDINAQIETNRNRLTLACQGAEEVYSLLDKRAVNEHRAKTGLTP 536
 QY 241 LMIASKEGHEIVODLDACTVYNIP--DRSGDTVLIGAVRGHVEIVALLQKYADIDI 298
 537 LMEASGTYAEVGRVLLDKGADVNAAPVPSSRDTALTIAADKHYFCELLHHRGHIDV 596
 QY 299 RGQDNKTALVWAVEKGNATVVRDILQCNPTETICTQDGETPLIKATKMRNIEVELLDK 358
 597 R-----NKKGVTPLMLASNGHFPVVDLYQA 623
 QY 359 GAKYSAVDKGDPTPLHAIIRGSRKLAELLBNPKXGRLLYRNKAGETPNYINDCSHQKS 418
 624 GADVDAADNRKTIPLMSAFKGVKYYQYLVK-----EYNQPSDIEC----- 666
 QY 419 ILTQIFGARHLSPETDGMGLGYDLYSSALADILSEPTMQPPICVGLVQWMSGKSFLLK 478
 667 -----MYTIA--TIIDKELL--KKCHQCVETIVAKKQDA-----AAKNNMSILK 709
 QY 479 KLEDE 483
 Db 710 ELDLF 714
 RESULT 7
 US-09-908-711-70
 ; Sequence 70, Application US/09908711

QY 383 KLAELLRN 391
Db 348 NCVOILLQH 356

RESULT 8

US-09-835-788A-18
; Sequence 18, Application US/09835788A
; Patent No. US20020077458A1
; GENERAL INFORMATION:
; APPLICANT: NI et al.
; TITLE OF INVENTION: Death Domain-Containing Receptor Polynucleotides, Polypeptides, and
; FILE REFERENCE: PTO18P1
; CURRENT FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: PCT/US00/28666
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/159,585
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 60/167,246
; PRIOR FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 18
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-835-788A-18

Query Match 4.2%; Score 376.5; DB 10; Length 285;
Best Local Similarity 36.7%; Pred. No. 3.8e-15;
Matches 98; Conservative 50; Mismatches 112; Indels 7; Gaps 5;

QY 43 MIAEOGNLEIVKELIKANGCNLEDD-NMTALISAKSGHVIYVELLKCYNLEHRD 101
Db 1 MEASAGHVEAVARVLDDHAGINTHSNEFKESALTCTGHDVDFLEAGADQHKHT 60
QY 102 MCGTALMMACYKGRITVDVELLSHGANGSVYTPILMAAGRGHADIYVHLLONG 161
Db 61 DEMTALMEACMDHVEAVARLLDLSGAQVNPADSPF-SPLTLAAGGHVLAALLIERG 119
QY 162 AKVNCSPKYGTTPLVMAARKGHECVYKHLAMGADV-OGCANSMTALIVAVNGYTOYV 220
Db 120 ANLEEVNDEGTTPLEBARSGHEMVALLLAOGANINAEETOEALTLACCGGSEVA 179
QY 221 KEILKRNPNVNLDDKGNLTALMTASKGHEIYVODLDAGTYNIPDRSGDYVLIGAVRG 280
Db 180 DFLIKAGADIEL--GCTPLMEASGEGHLEIVKYLILASGANVHATTATGDTALTYACEN 236
QY 281 GHVEIVARALLOKADIDIRGQDNKTAL 307
Db 237 GHTDVADVLLQAGADLD-KQEDMKTTIL 262

RESULT 9

US-09-972-115A-8
; Sequence 8, Application US/09972115A
; Publication No. US20030032769A1
; GENERAL INFORMATION:
; APPLICANT: Geron Corporation
; APPLICANT: Gregg, Morin B.
; APPLICANT: Walter, Funk D.
; APPLICANT: Mieczyslaw, Piatyszek A.
; TITLE OF INVENTION: A Second Mammalian Telomerase
; FILE REFERENCE: 080/003C
; CURRENT APPLICATION NUMBER: US/09/972,115A
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: US 60/128,577
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: US 60/129,123
; PRIOR FILING DATE: 1999-04-13
; NUMBER OF SEQ ID NOS: 64

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 8
; LENGTH: 1327
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-972-115A-8

Query Match 4.2%; Score 376.5; DB 9; Length 1327;
Best Local Similarity 22.3%; Pred. No. 3.6e-14;
Matches 160; Conservative 89; Mismatches 220; Indels 247; Gaps 17;

QY 14 EENINPRLKLLKCKDVERNECGOTPLMIAEOGNLEIVKELIKANGCNLEDDNKT 73
Db 348 EERLMLLPLPNVNCASDGRK--STPLHLAGYVRVYDILLDGHADVAKRGGLV 404
QY 74 ALISAKSGHVIHVEELIKGVNLEHRDGMGTALMACYKGRITVDVELLSHGANGSV- 132
Db 405 PLHNACSYGHEYTELLKRGACVNMMDLMQFTPLHEASKNRVEVCSLLSHGADPTLV 464
QY 133 -----T 133
Db 465 NCHGSAVDMPPELRERLTYEFKGSLLQAREADLAKVKTLALETINFKOPSHET 524
QY 134 GLQYV-----YPLIIMAGRGHADIYVHLLONGAKV 164
Db 525 ALHCAYASLHPRKKQYVELLKRKANVNEKNKDMFTPLHVAABERANDVMEVLRHAKAM 584
QY 165 NCSDKYGTTPLVMAARKGLECVKHLAMGADVDEGANSMTA----- 207
Db 585 NALDTLGTALHRAALAGHLQTCRLLSYSDPSIISLOGFTAAQMGNEAVQILSESTP 644
QY 208 -----LIYAVNGYTOYVKEILKRNPNVNLDDKGS--NTALMISAKSGHEIYVDL 256
Db 645 IRSDVDYRLLESKAGDLETVKOLCS--SQVNCBDEBRHSTPHFAGYNNRVSVEYL 703
QY 257 LDAGTYNIPDRSGDTVLIGAVRGHVEIVARLLQYADIDIRGQDNKTALYMAVERG-- 314
Db 704 LHHGADVHAKDQGLVPLHNACSYGHEIYVHLLSHGANGSVNADLMTKTPLEHAAKAKY 763
QY 315 -----NATVRDIIQ-----C 325
Db 764 EICKLLKRGADPTKRNKNDGNTPLDLYVEGDIIDILLKGAALLDAKCKGLARVQKLC 823
QY 326 NPDEICTKD--GEPPLIKATKMRNIEVELLDKAKVSYVDKSGPTPLH----- 374
Db 824 TRENINCRTOGRNTPPLHLAGYNNLEVAELLEHGADVNAQDGGILPLHNAASIGHV 883
QY 375 -----IAIRGRSRLAELLRNPKDGRILYRPNKAGETP 408
Db 884 DIAALLIKYNCVNAIDRWAFPLHEAAQKGT-OLCALLAHGADPTM--KNOEGQTP 939
QY 409 YN-----IDCSHOKSILT-----OIFGRHLSPETGDMIGLYLSALADI 451
Db 940 LDLATADIRALLIDAMPPEALPTCFKQATVVSASLISPASTPS-----CLSAASIDN 994
QY 452 LSEPTMOPPICGVIAOWSGSKSFLKK-----LEDEMTKFAQQOIEPLFQ 497
Db 995 LTGPLAE--LAVGASNMGDAAGATERKEGEVAGIDNMNISQFLASIGLEHRLDIFE 1048

RESULT 10

US-09-841-835-2
; Sequence 2, Application US/09841835
; Patent No. US20020076795A1
; GENERAL INFORMATION:
; APPLICANT: de Lange, Titia
; APPLICANT: Smith, Susan
; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue, 4th Floor

Db 465 NCHGSAVDMAPPELRELYEFGHSLLOAAREADLAKVKKTLAETINFKOPQSHET 524
QY 134 GLOYSV-----YPIIWAARGHADIVHLLONGAKV 164
Db 525 ALHCAVASLHPKKQYOTELLRRKANVNEKNDEMTPLVVAABRAHNDVMEVLLHKGAKM 584
QY 165 NCSDKRTGTPYWAARKGHEVCYKHLAMGADVDGANSMTA-----207
Db 585 NALDTLGGOTLHMAALAGHLQOTCRLLSYGSDPSIISLOGFTAAQNGNEAVOQILSESTP 644
QY 208 -----LIYAVKGYTSQYKELLKRPNVNLTDKOG--NFALMISKEGHTIYODL 256
Db 645 IRTSDVDYRLLEASRAGDEITYKQLOS--SONVNCBDEGRHSTPLHFAAGYRVASVVEL 703
QY 257 LBAQTYVNIPIRSGDVTLLIGAVRGHVEIVRALLOKYADIDIRGODNKTALYAVEK-- 314
Db 704 LHHGADVNAKDKGLVPLHNAACSYGHEVAELLVHGAASVNVADLMKFTPLHMAAKKY 763
QY 315 -----NATMVRDILQ-----C 325
Db 764 EICKLLKHGADPTKKNRDNTPLDLYEGDPTDIDLLKGDALLDAAKKGLARVOKLC 823
QY 326 NPDEICTKD---GETPLIKATKMRNIEVEYELLIDKGAKVSADVCKGDPPLHIAIR----- 378
Db 824 TPENINCRPTOGNRNTPPLHLAGYNNLEVAETLLEHGADVNAODKGLIPLHNAASYGCG 883
QY 379 -GRSKRLAELLRNPRKGRLLYRPNAKGETPYNIDC 413
Db 884 IARVOKLC-----TPENINC 898

RESULT 12

US-09-972-115A-6
; Sequence 6, Application US/09972115A
; Publication No. US20030032769A1
; GENERAL INFORMATION:
; APPLICANT: Geron Corporation
; APPLICANT: Gregg, Morin B.
; APPLICANT: Walter, Funk D.
; APPLICANT: Mieczyslaw, Piatyszek A.
; TITLE OF INVENTION: A Second Mammalian Telomerase
; FILE REFERENCE: 080/003C
; CURRENT APPLICATION NUMBER: US/09/972,115A
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: US 60/128,577
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: US 60/129,123
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 1166
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-972-115A-6

Query Match 4.2%; Score 370; DB 9; Length 1166;
Best Local Similarity 25.6%; Pred. No. 7.4e-14;

Matches 143; Conservative 61; Mismatches 169; Indels 186; Gaps 14;

QY 19 PALKALLEKCKD-----VDERNECGO--TPIMIAEOGNEIYKELIKNGA 62
Db 23 PAARELFEACRNGDVERVKRLVTPPEKVNSRDTAGRKSTPLHFAAGFRKRVYVELLONGA 82
QY 63 NCNLEDLDNWTALISAKKEGHVHVEELKCGVNLHRDMGWTALMAGCYKGRDVEL 122
Db 83 NVQARDGGGLIPLHNAACSGHAEVYVNLHRRGADPNARDNMNYPPLHMAIKGIDVCIY 142
QY 123 LLSHGAMPV-----TGLQY 137
Db 143 LLOHGAEPTIRNTDGRALDLADPSAKAVLTGEYKDELLESARSGNEKMMALLTPLNV 202
QY 138 SVY-----PIIWAARGHADIVHLLONGAKVNCSDK-----YG----- 171

Db 203 NCHASDGRKSTPLHLAGYNNKRYTOVLLDHGADVNAKDKGDLVPLHNAASYGHEYTEL 262
QY 172 -----TPYWAARKGHEVCYKHLAMGADVDGANSMTALIVA-----211
Db 263 LVKHGACVNAAMDLOFTPLHNAASKNREVCSSLLSYGADPTLLNCHKSAIDLAAPPOL 322
QY 212 -----YKG-GYTSQYKE-----222
Db 323 KERATYEFKHSLLQOAREADVTYRIKHLSEMYNEKHPOTHEALTHCAASYPKROI 382
QY 223 ---IKRNPVNLTDKDGNTALMASKEGHTIYODLLDAGTYVNIPIRSGDVTLLIGAVR 279
Db 383 CELLRRGANINTEKTEFLTPHNAASKANDVEYVYKHAAYNALDNIGTSLRAAY 442
QY 280 GGHYEIVRALLQKADIDIGODNKTALYAVEK-----GNATMVRDILQCNP- 327
Db 443 CGHLQOTRLLSYGCDPNIIISLOGFTALQNGNEVOLLQEGISLGNSEADROLLEAKA 502
QY 328 -DTE-----ICT-----KDEG---PPLIKATKMRNIEVEYELLIDKGAVSADVCKGDP 372
Db 503 GDEVTYKLLCTVOSVNCRDIEGROSTPLHFAAGYRVASVVELLHGHADVNAKDKGLV 562
QY 373 LHAIIRGRSKRLAELLRN 391
Db 563 LHNACSYGHEVAELLVKH 581

RESULT 13

US-10-163-587A-15
; Sequence 15, Application US/10163587A
; Publication No. US20030096263A1
; GENERAL INFORMATION:
; APPLICANT: Oliveira, Marcos
; TITLE OF INVENTION: SELECTIVE PARP-1 TARGETING FOR DESIGNING CHEMO/RADIO SENSITIZI
; FILE REFERENCE: 50229-306
; CURRENT APPLICATION NUMBER: US/10/163,587A
; PRIOR FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: 60/296,110
; PRIOR FILING DATE: 2001-06-07
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 1166
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-163-587A-15

Query Match 4.2%; Score 370; DB 9; Length 1166;
Best Local Similarity 25.6%; Pred. No. 7.4e-14;

Matches 143; Conservative 61; Mismatches 169; Indels 186; Gaps 14;

QY 19 PALKALLEKCKD-----VDERNECGO--TPIMIAEOGNEIYKELIKNGA 62
Db 23 PAARELFEACRNGDVERVKRLVTPPEKVNSRDTAGRKSTPLHFAAGFRKRVYVELLONGA 82
QY 63 NCNLEDLDNWTALISAKKEGHVHVEELKCGVNLHRDMGWTALMAGCYKGRDVEL 122
Db 83 NVQARDGGGLIPLHNAACSGHAEVYVNLHRRGADPNARDNMNYPPLHMAIKGIDVCIY 142
QY 123 LLSHGAMPV-----TGLQY 137
Db 143 LLOHGAEPTIRNTDGRALDLADPSAKAVLTGEYKDELLESARSGNEKMMALLTPLNV 202
QY 138 SVY-----PIIWAARGHADIVHLLONGAKVNCSDK-----YG----- 171
Db 203 NCHASDGRKSTPLHLAGYNNKRYTOVLLDHGADVNAKDKGDLVPLHNAASYGHEYTEL 262
QY 172 -----TPYWAARKGHEVCYKHLAMGADVDGANSMTALIVA-----211
Db 263 LVKHGACVNAAMDLOFTPLHNAASKNREVCSSLLSYGADPTLLNCHKSAIDLAAPPOL 322
QY 212 -----YKG-GYTSQYKE-----222

```

Db      323 KERLAEEFGHSLLOAAREADVTRIKKHLSLEWVNFKEHPOTHEHTALHCAASPYPRKQI 382
OY      223 ---ILKNPVNLTDKDGNALMIASKEGHEIYODLLDAGTYVNIIPDRSGDTVLIGAVR 279
Db      383 CELLRRGANINNETKEFLPLHVASKEANDVEVYVVKHAEVNMALDNGQSLHRAAY 442
OY      280 GGHVEIYRALLQYADIDIRGQDKTALYNAVEK-----GNATVRILOCPN- 327
Db      443 CGHLOTCRLLSYGCDENIISLOGFTALQMGNEVQOLLOEGISLSEADROLLEAKA 502
OY      328 -DPE-----ICD-----KDE-----PLIKATKMRNIEVELLDKGAQVSAVDKGDTP 372
Db      503 GGVETVKKLCTVOSVNCRDIEGROSTPLHRAAGINRVSYVEYLLQHGADYHAKDGSLVP 562
OY      373 LHAIIRGRSRKLAELLRN 391
Db      563 LHNACSYGHVEVALLVKH 581

```

RESULT 14

```

US-09-509-196A-2
; Sequence 2, Application US/09509196A
; Patent No. US20020037582A1
; GENERAL INFORMATION:
; APPLICANT: DALY, Roger J.
; APPLICANT: SUTHERLAND, Robert L.
; TITLE OF INVENTION: A Potential Effector for the Gb7 Family of Signalling
; FILE REFERENCE: 1871-129
; CURRENT APPLICATION NUMBER: US/09/509,196A
; PRIOR FILING DATE: 2000-03-23
; PRIOR APPLICATION NUMBER: P09368
; PRIOR FILING DATE: 1997-09-23
; PRIOR APPLICATION NUMBER: PCT AU98/00795
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1074
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-509-196A-2

```

```

Query Match      4.0%; Score 357; DB 10; Length 1074;
Best Local Similarity 23.9%; Pred. No. 4e-13;
Matches 150; Conservative 77; Mismatches 211; Indels 190; Gaps 13;
OY      14 EENIPALKALLECKDVERNEGQTPMTAEOGNLEIYKELIKNGANCNLEDDNMWT 73
Db      98 EKKMALLTPLNVNCHADGRK---STPLHLAAGYNRKIVQOLLQHRDVHAKDGDLV 154
OY      74 ALISAKEGHVAIYELLKCGVNIENRDMGNTALMACYGRDVLVELLSHGANSPT 133
Db      155 PLHNACSYGHEVEYELVKGCGVNAWMDLQFTPLHEAASKRNVESCLLSYGADPTLL 214
OY      134 G-----LOYS----- 138
Db      215 NCKNKSALDAPLPOLKERLAEFGKSHLLOAAREADVTRIKKHLSLEWVNFKEHPOTHE 274
OY      139 -----YPLIIMAGRGHADIVHLLONGAKV 164
Db      275 ALHCAASPYPRKQICELLRRKANINNETKEFLPLHVASEKHNDVVEVYVVKHAEKV 334
OY      165 NCSDKYGTTPVMAARKGLECVKHLAMGADV-----QEG-- 201
Db      335 NALDNLGQSLHRAAYCGHLOTCRLLSYGCDENIISLOGFTALQMGNEVQOLLOEGIS 394
OY      202 ---ANSMPTLIYAVVGTSYKE----- 224
Db      395 LENSEADROLLEAKAGDVEYVKKLCTVOSVNCRDIEGROSTPLHRAAGINRVSYVEYLL 454
OY      225 KNPVNVNLTDKDGNALMIASKEGHEIYODLLDAGTYVNIIPDRSGDTVLIGAVRGHVE 284

```

```

Db      455 QHGADYHAKDGSLVPLHNACSYGHVEVALLVKHGAAYVNAWDLKFTPLHEAASKGYE 514
OY      285 IYRALLQYADIDIRGQDKTALYNAVEKGNATVRILO----- 324
Db      515 ICKLLQHGADYHAKDGSLVPLHNACSYGHVEVALLVKHGAAYVNAWDLKFTPLHEAASKGYE 572
OY      325 CNPDPEICCKDE-----PLIKATKMRNIEVELLDKGAQVSAVDKGDTPPLHAIIRGR 380
Db      573 SSGPNVNC-RDYGGRHSTPLHLAAGYNLEVAEYLLQHGADYHAKDGSLVPLHNACSYG 631
OY      381 SRKLAELLRNPKDGRILYPRNKAGETPYNIDCSHOKSILIQIFGARHLSPT-----ET 434
Db      632 HVDVAALLI---KYNASLNATDKMAFTPLHEAAGRGTOQLALLAHGADYHAKDGSLV 688
OY      435 DGDMLGYDLYSSALADILSEPTMOPIC 462
Db      689 PLDLVSAVDVSAALLTAA--PPSALPSC 714

```

RESULT 15

```

US-10-164-080-7
; Sequence 7, Application US/10164080
; Publication No. US2003008741A1
; GENERAL INFORMATION:
; APPLICANT: BIRD, Timothy, A.
; APPLICANT: HOLLAND, Pamela, M.
; APPLICANT: PESCHON, Jacques, J.
; TITLE OF INVENTION: DEATH ASSOCIATED KINASE CONTAINING ANKYRIN REPEATS (DAKAR) AND
; FILE REFERENCE: 3280-B
; CURRENT APPLICATION NUMBER: US/10/164,080
; PRIOR FILING DATE: 2002-06-04
; PRIOR APPLICATION NUMBER: 60/295,959
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 60/334,362
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 784
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-164-080-7

```

```

Query Match      3.98%; Score 348.5; DB 9; Length 784;
Best Local Similarity 27.7%; Pred. No. 8.3e-13;
Matches 99; Conservative 64; Mismatches 126; Indels 69; Gaps 5;
OY      35 NECGQTPMTAEOGNLEIYKELIKNGANCNLEDDNMWTALISAKEGHVAIYELLKCG 94
Db      468 NRGSTPLHMAVERVAGVVELLAKISYNANDEQMTLHRAQNGDSSSTRLLLEN 527
OY      95 VNEHRDMGNTALMACYGRDVLVELLSHGANSPTGLQYVYPIIIMAGRGHADIV 154
Db      528 ASVNEVDGEFTPLHVAACQNGQENIVIRLLRGVDSLOQ-KDAMLPRLHYAAQGHLPY 586
OY      155 HLL-ONGAKVNCSDKGTTPVMAARKGLECVKHLAMGADVDOGANSMTALYAVK 213
Db      587 KILAKQGVSVNAQTOIDGRRPLHLAAGRGHVARILIDCSVNNCSLLAOTPLHNAE 646
OY      214 GGYTQSVKEILKRPVNVNLTDKDGNALMIASKEGHEIYODLLDAGTYVNIIPDRSGDTV 273
Db      647 TGHSTARLLHHRAGAEAMTSOYTLALHLA----- 678
OY      274 LIGAVRGHVEIYRALLQYADIDIRGQDKTALYNAVEKGNATVRILOCPNDEICT 333
Db      679 -----RNGHLATVLLVEERADYLABRPLQOTLHLAAAGHSEVEYELV----- 723
OY      334 KDGETPLIKATKMRNIEVELLDKGAQVSAVDKGDTPPLHAIIRGRSRKLAELLRN 391
Db      724 -----SADVIDLF-----DEQGLSALHLAAGHQAQTVETLLRH 757

```

Tue Jul 1 15:13:30 2003

us-10-021-571-4.rapb

Page 10

Search completed: July 1, 2003, 14:51:10
Job time : 51 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2003, 14:34:05 ; Search time 19.5 Seconds
(without alignments)
2587.708 Million cell updates/sec

Title: US-10-021-571-4

Perfect score: 8884
Sequence: 1 MSYLISQSYNTVEENIPA.....LHMAASSESTGGEHRESL 1715

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PCUS.COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	530.5	6.0	1839	US-09-172-977-4	Sequence 4, Appl1
2	493.5	5.6	843	US-09-172-977-3	Sequence 3, Appl1
3	480.5	5.4	1745	US-09-031-485-33	Sequence 33, Appl1
4	480.5	5.4	1745	US-08-847-429A-33	Sequence 33, Appl1
5	480.5	5.4	1745	US-09-065-474-33	Sequence 33, Appl1
6	480.5	5.4	1745	US-09-557-034-33	Sequence 33, Appl1
7	438	4.6	1088	US-09-082-059-2	Sequence 2, Appl1
8	409.5	4.9	1423	US-08-810-712-10	Sequence 10, Appl1
9	376.5	4.2	1327	US-09-196-387-2	Sequence 2, Appl1
10	374	4.2	348	US-09-031-485-28	Sequence 28, Appl1
11	374	4.2	348	US-08-847-429A-28	Sequence 28, Appl1
12	374	4.2	348	US-09-065-474-28	Sequence 28, Appl1
13	374	4.2	348	US-09-557-034-28	Sequence 28, Appl1
14	370.5	4.1	949	US-09-196-387-10	Sequence 10, Appl1
15	363	4.2	1166	US-09-350-982C-5	Sequence 5, Appl1
16	356.5	4.0	352	US-09-065-474-139	Sequence 139, App
17	356.5	4.0	352	US-09-557-034-139	Sequence 139, App
18	353	4.0	302	US-09-031-485-38	Sequence 38, Appl1
19	353	4.0	302	US-08-847-429A-38	Sequence 38, Appl1
20	353	4.0	302	US-09-065-474-38	Sequence 38, Appl1
21	353	4.0	302	US-09-557-034-38	Sequence 38, Appl1
22	353	4.0	303	US-09-031-485-23	Sequence 23, Appl1
23	353	4.0	303	US-08-847-429A-23	Sequence 23, Appl1
24	353	4.0	303	US-09-065-474-23	Sequence 23, Appl1
25	353	4.0	303	US-09-557-034-23	Sequence 23, Appl1
26	353	4.0	741	US-08-436-771-4	Sequence 4, Appl1
27	353	4.0	741	US-08-434-998-4	Sequence 4, Appl1

28	353	4.0	741	US-08-487-797-4	Sequence 4, Appl1
29	353	4.0	741	PCT-US95-02058-4	Sequence 4, Appl1
30	348	3.9	679	US-08-462-481-4	Sequence 4, Appl1
31	348	3.9	679	US-08-436-771-6	Sequence 6, Appl1
32	348	3.9	679	US-08-434-998-6	Sequence 6, Appl1
33	348	3.9	679	US-08-487-797-6	Sequence 6, Appl1
34	348	3.9	679	US-08-701-005A-4	Sequence 4, Appl1
35	348	3.9	679	US-08-479-895-4	Sequence 4, Appl1
36	348	3.9	679	PCT-US95-02058-6	Sequence 6, Appl1
37	344	3.9	679	US-08-943-956A-4	Sequence 4, Appl1
38	344	3.9	787	US-09-188-930-334	Sequence 334, App
39	343	3.9	741	US-08-462-481-2	Sequence 2, Appl1
40	343	3.9	741	US-08-436-771-2	Sequence 2, Appl1
41	343	3.9	741	US-08-434-998-2	Sequence 2, Appl1
42	343	3.9	741	US-08-487-797-2	Sequence 2, Appl1
43	343	3.9	741	US-08-701-005A-2	Sequence 2, Appl1
44	343	3.9	741	US-08-479-895-2	Sequence 2, Appl1
45	343	3.9	741	PCT-US95-02058-2	Sequence 2, Appl1

ALIGNMENTS

RESULT 1					
US-09-172-977-4					
; Sequence 4, Application US/09172977					
; Patent No. 5989863					
; GENERAL INFORMATION:					
; APPLICANT: Tang, Y. Tom					
; APPLICANT: Guegler, Karl J.					
; APPLICANT: Corley, Neil C.					
; APPLICANT: Yue, Henry					
; TITLE OF INVENTION: HUMAN ANKYRIN FAMILY PROTEIN					
; FILE REFERENCE: FF-0615 US					
; CURRENT APPLICATION NUMBER: US/09/172, 977					
; CURRENT FILING DATE: 1998-10-14					
; NUMBER OF SEQ ID NOS: 4					
; SOFTWARE: PERL Program					
; SEQ ID NO 4					
; LENGTH: 1839					
; TYPE: PRT					
; ORGANISM: Homo sapiens					
; FEATURE: -					
; OTHER INFORMATION: 929491					
US-09-172-977-4					
Query Match					
Best Local Similarity 20.2%; Score 530.5; DB 2; Length 1839;					
Matches 426; Conservative 263; Mismatches 654; Indels 763; Gaps 81;					
QY	14	EEENIPALKALKECKDVEDERNECGOTPLMAEONGLEIVELKINGANCLEIDDMNT	73		
DB	73	KEHGVGLVOLLRLRGSSVDSATKKGNTALHISLAQAEVYKLVKGNINAAQSGNFT	132		
QY	74	ALISAKKEGVHIVLEELKGVN-----LEHRDNG---	103		
DB	133	PLYMAQENHIDVYKLENGANQSTATEDGFTPLVALQGHNAQVAILLENDTKGVR	192		
QY	104	-----GWTALMWACYKGTVDVVELL	123		
DB	193	LPALHAARKDDTKSALLLQNDHNDVOSKMNVRFTSGFTPLIAHVGWVAVATLL	252		
QY	124	LSHGANSVYGLQSYVPIITMAARGHADIVHLLONGAKVNCSDRYGTTPLVMAARKS-	182		
DB	253	LNKGAADVFTA-RNGITPLVAVAKRGTNNVKLLDRGQIDAKTDIGLPLHCAARSH	311		
QY	183	-----HLECYKHLANGADVDQEGANSMTALIV	210		
DB	312	DQVVELLEKGAFLARTKNGLSPLHMAAGDHEVCVKHLLDKAKAVDVTLDYTLAHV	371		
QY	211	AVKGGYQSVKEIKRPNVNLDDKQGNTPALMASKE-----	247		
DB	372	AAHGHYRVTKLLDKRANPNARALNGFTPLHACKKNRIKIKWELLVYKGASIOATTESG	431		

```

OY 248 -----GTEIVODLLDAGTVNIPDSRSDTVLIGAVGSHVEIVRALLDOKYADID 297
Db 432 LRPVHAAMFGLNLTIVLLLLONGASPDVTNIRGETALMMARAGVEYVRCILRRGALVYD 491
OY 298 IAGONKTAIYAVEKGNATWVRLQCNPDTEICTKDEPPLIKATMRNIEVEVELLD 357
Db 492 AARBEQTPPLIASRLGKTEIYQILLQHMARDATNGYTPPLHISAREGOVDVAVLE 551
OY 358 KCAKVSAYDKKGDPLHAIAGRSKLAELLR-----NKPD 394
Db 552 AGAASLTKKGFTHVAAKIGSLDVAKLLQRAAADSAGKGLPLHVAHAHYDMOKV 611
OY 395 GRLI-----YRPNKAGETPNYIDC-SHOKSILTOI--FGAR-----HLSP 432
Db 612 ALLLEKASPAHATKNGYTPPLHIAKKNQOIASTLLNAGAEINVTYKOGVTPPLHLSAQ 671
OY 433 EDDGML-----GYDLSSALADLSEPTMOPPCVGLYQWQSGSKFLKLKLEDEM-- 484
Db 672 BGTHTVTLTLDKGNIMHSTKSGLTS-----LHLAAO-----EDKYNV 710
OY 485 -----KTPAGQOIEPLFOFSLIYELTLLC--GGIGL--PFTVHPMLGIAVSLFLAL 536
Db 711 ADILKAGADDADHAKLGYTPPLIY-----ACHYGVKKNVNFLLQGANVNAKTKNGYTP 765
OY 537 -----LYFFIYIYFGRRGEGSNWMAVLSRLARHIGYELL--IKLM----- 579
Db 766 HQAAOQGHTHIIVLQHGAKPNATYANGNTALA--IAKRLGYISVDTLKVTEVEVYTT 823
OY 580 -----FVNPELPEQTKALP--VRPL-F 600
Db 824 TTTITEKKNVPEMTETVLDVSDDEGDTMTGCGEYLRDEDLKEJDDSLPSSQPLDG 883
OY 601 TDYNLSSVGET--SL-----AEMTILSDACERE 629
Db 884 MNYLRYSLGGRSRLBSFSDRSHTLSHASYLRDASVMDSVIYPSHQVSTLAKEMERN 943
OY 630 FGLATLRFYRKFKEDYOGKKKKKTCCLPSFYFLP-----IIGCTISGILLAFRV 683
Db 944 SYRLS-----WGTEINDNALSSPSIHSGFLVIFMVARGARGCRHNLRII--I 993
OY 684 DPKHLYNAVLIASVYGLAFVLCRTMQLVDSLNSOKRL-----HNAASKL 734
Db 994 PPRCTAPL-----RYTCHL-----VKHRIATMPVVEGGLASRL 1030
OY 735 HKLSEG-----FMVKLCEVELMAMARTIDSFTON-----QTRLVYIIDG 776
Db 1031 IEVPSGAGFLPYIVEIPHPAALRGKERELVLRSENGDSMKHEFCDYDEDELINELNG 1090
OY 777 LDAGEOKVLOMLDTVAVLPSKGFIAIFASDPH--IIKAIQNLNSVLRDSNINGHY 834
Db 1091 MD-----EVLDSPEDELEKRICTIITRDFQYFAVVSRIKQ-----DSNLGPE- 1134
OY 835 MRNIVHLPVFNLSKGLSNARKFLVTSATNGDVPCSDTGTGQIADRRVSONSGEMK-- 892
Db 1135 -----GGVLSSTVVOQVAVPEGALTKRI--RVGIQ-----APMISELYKTI 1176
OY 893 LGSKTALNRDRTY--RRQOMORTTROMSDELTKLVLTEWFSDISQTMRLINIVSVT 950
Db 1177 LGNATESPPLYTLBRKRRKHKPITMTIPVPKASSDVMLNGFGGDAP--TLRL--CSIT 1232
OY 951 GRLRANOISFNMRLASWINLT--EQWPYRTS-----WLIILETBEDIPOMLTKT 1002
Db 1233 GGTPA-----QWEDITGITPLTFVNECVSFTTNVSAARFILI--DCROIQESVTFASQ 1283
OY 1003 YERISKNIPTTKDVEPLEITDGRNEFVLSKTPPLVARDVYVLPCTVNLDPKL-- 1059
Db 1284 VYREITGVY-----MAKFVFAKSHD-----IEARLRCCMDDKVDKTI 1324
OY 1060 ---REITADYARAEOISIGLAVPPLLHEGPPRAVSGYSQPPVCSSTSFNGPPAGV 1116
Db 1325 LEODENFAEVARSHD-----VEVLEGRPIYVDCF-----GNL 1356

```

```

OY 1117 V-----SPQHSYYSGMTGPHPEFYNRGSGPAPGVVLLNSLVDAVCEKIKQIEGDS 1172
Db 1357 VPLTKSQHHIFSEFAFKENRLPLFYKVRDTPDPCGRUSFM-----KEPSTRGLVHQ 1410
OY 1173 MLPOYCTTIKANINGVLAQCNID-----ELKEMNNFGDHLFRSTVL 1218
Db 1411 -----AICMLNTLPIYTKESSDQEQEEDIMT----- 1439
OY 1219 EMNASHVYVPEPPRELSESSGPPADHPGEPARASINELPHTELSCQTPYTLNFSFEELN 1278
Db 1440 -----SEKNPODEQERIEE-----RLATYADHLAGFSEWELARE--LDFTEEQTH 1481
OY 1279 TLGDEGAPRSHNLSMQSO-----TRRPSLSLNSODSSEISKILPDK 1322
Db 1482 QIRIE-----NPNSLOQSOYLKIMLERQGNATIDNLVCECLKIRMDIYHLMETNERP 1537
OY 1323 VQAEYRDATREYIAQMSQLEGFGSTTIGR--SSPHSTYWGQSSSGSIHSLNEQEK 1380
Db 1538 LOERISHSYAE--IEQTITLDHSEGFSLQELCTAOHK-----QKEBOAVS 1582
OY 1381 KDEPPEPDGGRKSFLLKRGVDIDSSSGYSTNDASPLDITTED-----EK 1426
Db 1583 KESE-----TCDHPPI--VSEBDSISVGTSTQDGVPKTE 1614
OY 1427 SDQSGSKLLPGKKSSESSLFOTDLKLGSGLYOKIPSPDED--ESQTERSDMTPLKDK 1485
Db 1615 GDSSTALFPQTHKEQ--VOODF--SG-KMDLPRESSLEFQOEYFTVTPGETISE 1665
OY 1486 DRKAEKVERVYKSPSHSAPITFTIKAEYLSDALDKKSDSGVRSSESPHNSLHN 1545
Db 1666 TORA--MIVPSSPKTPE----- 1681
OY 1546 EVADDQLEKANILEEDDSHGKRGIPHSLSGLDOPITARSICSE--DKKSPSECSLI 1603
Db 1682 EVSTPADEER--LYIQTPSSERG--SPIIQEPPESEHREESSPKRTSLV 1729
OY 1604 ASSPEENWP-----ACQKAYNLNR--TESYTLNNSAPANRANONFDEMGIIR- 1651
Db 1730 IYVSADNOPETCERLDEDAFAFEKDDMPEIPEYVT-----EEBYIDH 1773
OY 1652 -----TSQVILRSSS-----PNTTIONMLSMTHKRSQBSYRL 1689
Db 1774 GHTVAKVTKRIITIRYVSEGETEKEIIVQGMPOEVNIEEDGYSKYIKRVYLSKSDTEQ 1833
OY 1690 SKDPPE 1695
Db 1834 SEDNNE 1839

RESULT 2
US-09-172-977-3
; Sequence 3, Application US/09172977
; Patent No. 5969863
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN ANKYRIN FAMILY PROTEIN
; FILE REFERENCE: PF-0615 US
; CURRENT APPLICATION NUMBER: US/09/172,977
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PERL Program
; SEQ ID NO 3
; LENGTH: 843
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: g1841966
US-09-172-977-3

Query Match 5.6%; Score 493.5; DB 2; Length 843;

```



```

QY 392 -----PKDRLLYRNKAGETPYNIDCSHOKSILNIOIFGARHLSPETDGMIG 440
DB 686 QEDRVSAEELVKENAAIDPKTKAGYTPLVHACH-----IG-----QINMVR 727
QY 441 YDLSSALADILSEFTMPPICVGLIYQWSSKSLKLEDEKMTFAGQOIEPLPQPSW 500
DB 728 FLIEHGANVSITRASYR-----LHQAAQGHNSVR-----760
QY 501 LIVETLLLCGLIGLFAFVHPNGLIYVLSLFLALIFPIVIFGGRGESMNAW 560
DB 761 -----YLEHGASPNVHTSGOTPLS-----781
QY 561 LSTRARIHYLELLIKLMFVNPELPOTTKALPVRE-----LEFDYNRLSSV 609
DB 782 -----IARLGYVSVEALKTTEFTVITETTVTEERYKPNPEAMETMSD-----SED 833
QY 610 GGETSLAMIMTSLSPACEREGFL--ATRLFRVKTEDTQCK-----650
DB 834 EGEDNQTANAHADFSLSLTKGLHDSGVHLHATEPTLSRSEVEGTDDLDALIRKA 893
QY 651 -----KWKTCCLPSFVIFLEI-----IGCIISGTLTALF 681
DB 894 QHEPITAMADPSIDASLPDNTVIMRTMQPSFLISFWVDARGAMGCRHSYRIIIP 953
QY 682 RVDRKHLTVNAVLSIASVGLAVLCRTMMOVLDSLN-----SQRRLHNAASKLHL 737
DB 954 RKAQOPTRV-----TCR--YLGCKDLAPHPPLSEGEAL--ASRILEM 991
QY 738 KSEG--FMKVLKCEVELMARMAKTIDSEFONOTRLVYII--DG-----LQACEDKV 765
DB 992 APHAKELGVITLEVPHASL-----KREKEIYILNSDDQHKEHOLEATE--DAV 1042
QY 786 LQMLD-----TVRLFSGPFIALFASDPHIIKAINONISVLRDS 827
DB 1043 QEVINESFDELSQDLDTLSRITRLTNDPFI--YEA-----VYIRVQEVHCVGPEG 1095
QY 828 NINCHDMRNTV--HLPFVLSRGLSNARKLVSATNGDVPCESTGICQDADRVSONS 886
DB 1096 GV-----LSSVVPVQVLPDGSIT--KIKVSVQAPVPOEIVTRLH-----1137
QY 887 LGEMTKLGSKTALNRDVT--RRROMORTITROMSFDLTK--LVTEDWFSISPQTM 941
DB 1138 -----GNRYVASPIYVTEPRRRKFKPITLICPLRQSNKMLVOYSQPOQEPPTLR 1190
QY 942 RLL-----NIVSVTGLIRANQISEN-----W-----DLASMT--N 971
DB 1191 LLSKGTGSSPAQMEDITGTTOLFTGEDVSTTVSARFWMDCOTPRDARMAQEVYN 1250
QY 972 LTEQWPTYSMLILETEGIDPOMTL-----KIT--YERISKNIPTTKDEYEL- 1019
DB 1251 EATAVPMYMAKFLI-FARTRPAGQLRFGMTDREDKYLEKOEKETE--IAKSKDEVILS 1308
QY 1020 -----LEIDGDIRNEFEVLSRRPVLVARD--VKVPLPCTVN--LDPKLR-----EITA 1064
DB 1309 GRHOFEFSGNL-----LPITKSGDQLSLYFLPQENRLAMVIRTRTNDETA 1358
QY 1065 DVNAARQISIGGLAVPLPLHGGPRRABGYSQPSVCSSTSFNPPRAGVGVSPPHS 1124
DB 1359 DGR-----IYMKERKLAENLPPOTP-----VCTLATITPEYTG--PEPMVS 1399
QY 1125 ---YYS--GMTGPHHFYNGSGPAPR--VVLNLSLWMDVCEKTK-----Q 1165
DB 1400 KILFYEASLITEKYVAFHETAPRDLPLAHVALLIGAMHRLARALEVRDIDIKOVRQ 1459
QY 1166 IEGLDQSMLEPYCTTTIKKANIN--GRVLAOCNIDELKEMNM-----F 1207
DB 1460 LVGLLEAVITLRIWIFLKKEDAPVALRSALORIGRDVVRREMDRAKLDGLBTPVSHS 1519
QY 1208 GDWHLFRSTVYLEM--RNAE-----SHVPRDPRRLSSSG--PAPHEPRARRASHNL 1257
DB 1520 GRSITISLLEAVAGRRRHAETVMAOQRLADPPFOOVGYNGTPEDEPREEQSFHEE 1579
QY 1258 PHTLESSQTPYTLNFSFEELNLTGLDEGARPHRSNLSWOSQTRRTPSLSLNSQDSSIEIS 1317

```

```

DB 1580 EBAVVS-----EIRIVKTE--RHVHDSNGPIYEERTITTTTIEDDAVNEE 1624
QY 1318 KLTQV-----QAEYRAYREYIAQMSQLEGPGSTTISGRSPHSTYMGQSSSGS 1370
DB 1625 EIVDKIYPLNEEQEKKDRVRVREVENNFQOE-----TSKEGFGCOTTHEKEDDGS 1679
QY 1371 IHSNLEQKKQSEPP--DGRKSFLLKRGVDYDSSGVSTNDASP-LDPTIED 1424
DB 1680 LKQTM-----KSHVHQIFPDGGETS-----ANETGLSSGDADTITMPTTKED 1722

RESULT 4
US-08-847-429A-33
; Sequence 33, Application US/08847429A
; Patent No. 5827692
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESS: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: Windows 95
; SOFTWARE: Wordperfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/847,429A
; FILING DATE: 24-APR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1745 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-847-429A-33

Query Match 5.4%; Score 480.5; DB 2; Length 1745;
Best Local Similarity 20.3%; Pred. No. 5,66-32;
Matches 353; Conservative 234; Mismatches 602; Indels 549; Gaps 67;

QY 14 EENPILAKLLECKVDNERNEGQTPPLMAAQMLEIYELKILKAGANCLELDN-- 71
DB 207 KKDPTKATLTLQNEBHNSDVTSKSGFTPLHAHYGNENVAQILLEKGANVNOARHNIS 266
QY 72 -----W-----TALISAKSGAHYIYEELKKG-----94
DB 267 PLHVATKGRKNTWSLLAHGAVIDCTROLITPLRHCASSGHDQVVDLLEKAPISAK 326
QY 95 -----VNLH-----EDMGWTALMW 110
DB 327 TKNGIAPLHMAQVDVTVYITPLHVAAGHVRVAKLLIDRNADPNBALNGFTPLHI 386
QY 111 ACYGRDQVVELLSHGANSVYGLQYSYVPIIAGRGADIVHLLDNGAVNCSDKY 170

```

```

Db 387 ACKNRKIKYELLKTHAIMEAT-TESG:SLPLHVAFMCAINIVYILLQOGANADAVTR 445
QY 171 GTPLVMAARKHLECYKHLILANGADYDOEGANSMTALIVAKGTTOSYKELKKNPV 230
Db 446 GETPLHAAANQTDIVRLVLRNGAQVDAARELQTPHLIASRLGNTDVIILLQANASP 505
QY 231 NLTKGKNALMAKSGHEIYQDILDDGYVNIIDRSGDPTLIGAVGHEIYRALL 290
Db 506 NAATRDLYTPHLIAEGOEVAAILMDHGTDLTKKGTPLHLIAKGNLPAVKSL 565
QY 291 OKYADDIRGODKKTALYMAVEKGNATMVNDILQCPDTEICTKDETPLIKATKRNIE 350
Db 566 ERGTPVDIEKNGVTPHLVAHAHNNDKVALLENGASAAKAKNVTPLHLIAKKNMD 625
QY 351 VELLIDKAKVAVDKGDTPLHLIARSGSRKLAEILLRN 391
Db 626 IASTLLHYKANANAESAGFTPLHAAOEGHREMAALLIENGAKVGAQANGLTPMHLCA 685
QY 392 -----PKDGLLYPNKAGETPNYIDCSHQSILTOJFGARHLSPETDDMLG 440
Db 686 QEDRVSAEELVKNENAIIDPKTAGITPLHVACH-----FG-----QINMVR 727
QY 441 YDIYSSALADILSEPTMQPICVGLXAWMGSGKSLKLEDEMKTFAGQIPLFOFSW 500
Db 728 FLIEHGARVSVITRASVTP-----LHOAQOGHNSVVR-----760
QY 501 LIYFLITLLCGGLLFAFTVHRLGIAVSLFLALTYIFYIYIRGGRREGESWMAWY 560
Db 761 -----YLLEHGASPNVHTSTGOTPLS-----781
QY 561 LSTRARHIGYLELLILKLMFVNPELPDEQTKALPVAF-----LFTDYNRLSV 609
Db 782 -----IARLGVSVAELAKTITETVTEETTYEENYKQNPENAMETFS-----SED 833
QY 610 GGETSLAEMATLSDACEREGFL--ATRLFRVFKTIDTQCK-----650
Db 834 EGEDNOITANAHADFESESLTKGLHDSGTGVLHATEPTLSRSEVEBGTDGDDALIRKA 893
QY 651 -----KWKTCCLPSFVIFLFI-----IGCIIISGTLIAIF 681
Db 894 OHEPTTAMADPSLDASLPDQNVITMRTMOPSLISMVADRGAMGCRHSVRIITPP 953
QY 682 RVDPRHLTVNAVLISIASVGLAFVLNCRWMOVLDSILN-----SQRKRLNASHLKL 737
Db 954 RKAQPTPV-----TCR--YLCKDKLAHPPLSEBAL--ASRLTEM 991
QY 738 KSEG--FMKVKCEVELMARMAKTIDSFTONQRLVYI--DG-----LDACEQDV 785
Db 992 APHGAKFLGPVILEVPHFASL-----RGEREIYILRSDDGQHKHEHOLEATE-DAV 1042
QY 786 LQMLD-----TVRYLFSKGPRIAFASDPHIIKAINONLSVLDS 827
Db 1043 QEVINESFDABELSOLDLHSTRITRILLNDPFA--YFA-----VTVRROEVHCVCPBG 1095
QY 828 NINCHDYARNIV-HLPEFLNSRGLSNARKFLVTSATNGVPCSDTGTIOEDADRVSNS 886
Db 1096 GV-----LLSVYVPHQAIFPDGSL--KTIVSVQAQPVQELVTRLH-----1137
QY 887 LGEMTKLGSKTALNRDITY--RRQOMOTTIRQMSFDLTK--LLVYEDWFSDISPOTMR 941
Db 1138 -----GNRVAVSPIVVEPRRRKFRHPIRLCIPLOSNNKMLTOYSOGPGEPPTLR 1190
QY 942 RLL-----NIVSVTGRLLRANQISFN-----W-----DRLASW-N 971
Db 1191 LLCSKTGSSSPAQWEDITGTTOITFTGADVFTTIVSARFLMOCQTPRDAARAQOEVYN 1250
QY 972 LTEOWPYRTSWLILYEETEGIPDOMTL-----KTI--YERISKNIPTKDOVEPL- 1019
Db 1251 EAIANVYAKFLI-FARTPFAEQGLRFLCMTDREDEKTLKOEERFIE-IAKSDOVEYLS 1308
QY 1020 -----LEIDGDIRNEVEVLSRTPVLVARD--VKVFLPCYV--LDBKLR-----EIIA 1064
Db 1309 GRHOFLEFSGNL-----LPTTKSGDOLSLVFLPQENRLAFVWKIRTHDNETAA 1358

```

```

QY 1065 DVRAAREQISIGGLAYPLPLHEGPPRAPSGVSGPPVSCSSSEFNGPAGVSPQBRSS 1124
Db 1339 DGR-----IYMEKPKLRANLPPQTP-----VCTIAITLPEYTG-----PEPVVS 1399
QY 1125 -----GMTPGPPHFNRRSGPAPG--VLLNSLNVAVCEKIK-----Q 1165
Db 1400 KLFYSSEASLPEKUYGAFHETAEPNLPLAHVALLIGADHRLARALEVPDIDIRQVHQ 1459
QY 1166 IEGLDQSMPLQYCTTIKKANIN--GVLACNIDELKEMNN-----F 1207
Db 1460 LVGLAEVITLRIMPLIKKEQATPVALRSALORIGDQVREMDRAEKLDLEGTPVSHIS 1519
QY 1208 GDMHLFSTVLEM-----RMAE--SHVPEDPFLSESSSG-PAPGEPARASHNEL 1257
Db 1520 GSTITSLSTLLEVAQDRRHAEVYTAQAORLAQEPFOOVGNGTGDPEDEKEOSFHEE 1579
QY 1258 PHTELSSQTPYTLNFSFEELNTLGLDEGAPRHSNLSWOSQTRRPSLSSLSNDSOSSIIS 1317
Db 1580 EEVAVS-----EITVVRTE--RHVHSENGPIVEERTITTYEDVAVNEE 1624
QY 1318 KLTDKV-----QAEYRAVREYIAQMSQLEGGGCTTISGRSSPHSTYTMGSSSGS 1370
Db 1625 EIVDKIVPLNEEEOEKMDRAREVEKMEPEQOE--TSKEGTFCQOTHEKEKDDGGS 1679
QY 1371 IHSNLEQEKGDSEKPP--DDGRKSFLMKRGDVIDYSSGVSNDASP-LDPTIIEED 1424
Db 1680 LKTTM-----KSHVNAQIFPDGEMS-----ANETGLSSDADPTIMPTTKED 1722

RESULT 5
US-09-065-474-33
; Sequence 33, Application US/09065474
; Patent No. 6063599
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 171
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESS: Hesk Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: Wordperfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/065,474
; FILING DATE: 24-Apr-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5-C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1745 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-065-474-33

```

Query Match 5.4%; Score 480.5; DB 3; Length 1745;
 Best Local Similarity 20.3%; Pred. No. 5 6e-32;
 Matches 353; Conservative 224; Mismatches 602; Indels 549; Gaps 67;

14 EEBNIPALKECKDNERNEGOTPLMTAEOGNEIYELIKNGANCLEJEDLN-- 71
 207 KDDKRAATLLQNEHNSDVTSGFTPLHIAHGNENVAQLLEKGANVYQARHNS 266
 72 -----W-----TALISKEGHVIVELLKCG----- 94
 267 PLHVATKGRIMVSLAHAGVADICRTDRDLTPLHCASRGHDVDLLEKGPISAK 326
 95 -----VNLH-----RDMGWTALMW 110
 327 TNGIAPLHMAQVDVYDYLPLHVAHAGHVAVKLLDRNADPARALNGTTPH 386
 111 ACYKRTDVELLSHGPNPSTVGIQYVPIYMAAGHADYVLLQNGAKYNSCKY 170
 387 ACKKRIKIVELLHYHAIEAT-DESGISPLHVAFMGAINIYVLLQGANMDVAVR 445
 171 GTPLVMAARKGLECVKHLAMGADVDQEGANSMTALIVAKGYTQSVKEILKRNPNV 230
 446 GETPLHIAARAQOTIVIVLVANGQVAAARELOTPLHIASRGNTDYLVLQANASP 505
 231 NLTDDGNTALMASKEGHTTEVDLDAGTYVNIIDRSSGTYVLGARGGHVETVAL 290
 506 NATRDLYPLHIAKEGQEEVAALIMDGTDKTLTKKGFPLHIAKYCNLPAKSL 565
 291 OKYADIDIRGODNKTALYMAVEKGNATVNRDILQCNPTETCTDGEPLIKAKRMNIE 350
 566 EGTGYVDIEGKNQVPLHVAHYNDRKALLLENGASAHAAANGYPLHIAAKQMD 625
 351 VELLIDGKAVYSAVDKGTPLHIAIRGSRKLAELLRN----- 391
 626 IASTLLHYKANANESKAGFTPLHIAQEGHREMAALLIENGAVGAQARGLTPMHLCA 685
 392 -----PKGRLLYPRNKAGETPNYNDCHOSITLQIFGARHLSFTEDGMLG 440
 666 QEDRVSAEELVKEAALDPKTKAGYTPPLHVAHCH-----FG-----QIMNVR 727
 441 YDVSSALADILSEPTMOPICVGLYAWGSGSKFLKLEDEKTFAGQOIEPLFQFSW 500
 728 FLIEHAGARVITRASYTP-----LHQAAQGHNSYVR----- 760
 501 LIVFTLLCGGLGILFAFYVHPMLGIAVLSLFTALLYFTFYVFGRRGESWMAWY 560
 761 -----YLHEGASPNVHTSGOTPLS----- 781
 561 LSTRLAHIGYELLKLMFVNPDELPEOTTKALPVR-----LTDYNRLSV 609
 782 -----IAERLGVSYVEALKTTETTYVITETTYTEERKKPONPEANMETWSD---SED 833
 610 GGSTSLAMIAVTLSDACEREGFL-ATRLFRVKTEDTQGGK----- 650
 834 EGEDNQITANAHADFESLTKGLHDSIGVHLIHATEPTLSRSEVEVGTGDALIRKA 893
 651 -----KWKTCCLPSFYVIFLT-----IGCIISGITLLAIF 681
 894 QHEPITAMADPSLDASLPNVNITMRTMOPSLISFMVARGAMGCRHSGRITIPP 953
 682 RVDKHLTVNAVLISTIASVGLAFVLCRTMOWLDLNLN-----SOKRKLHNAASKHL 737
 954 RKAQOPTRV-----TCR-YLCKDKLAHPPLSEGL---ASRLKM 991
 738 KSEG--FMKYKCEVELMARKATIDFTQNGTRLVII--DG-----LDACRODKV 785
 992 APHGAKFLGPIYLEVPHFASL-----RGREIEYILSRSDGQHHKEHQLATE-DAY 1042
 786 LQMD-----TVRLSKGPPIAFASDPHIIKAINQNLNLYROS 827
 1043 QEVYNESFDAEELSOLDLHTSRITRLITNDPFW-YFA-----VTVRQVQEHVCPREG 1095
 828 NINCHDYMRNIV-HLPEVLSRGLSNARKFLVTSATNGDVPCSDTJIGDADARRYSQNS 886

1096 GV-----ILSSVPPVQAFPPDGSILT---KTIKYSVAQAPPOBIYVRLH----- 1137
 887 LGEMTKLGSKTALNRDPTV--RRROMRTITROMSFDLTK---LVTEDWFDISPOTR 941
 1138 -----GNVAVSPITVPERRRKPKFTTLCIPLRPSNKMMLQYSQPOQEPPTIR 1190
 942 RLL-----NIVSYTGRLLRANOISEN-----W-----DLASWI-N 971
 1191 LLSCKTGSSPAQWEDITGTTLFTGEDVSFTTVASARFWMDCQTPDAARMAQEVYN 1250
 972 LTBQWPRYSMLIYLEETEIPDOMTL-----KIT-YERISKNIPTKTRDEPL- 1019
 1251 EATAVPMARFLLI-FARTTPAGQLRLFCMTDREDKTIKEKERTIE-TASKDVEVLS 1308
 1020 -----LEIDGDIRNEFEVLSRRPVLVARD--VKVFLPCTVN--LDPLR-----EITA 1064
 1309 GRHOFLEFSGNL-----LPITKSGDLSLYFLPQENRLAPMVKIRTHDNETA 1358
 1065 DVNARARQISIGLAVPLPLHGGPPRAPBSYQSPSVCSSTSFNCPFAGVYSPHSS 1124
 1359 DGR-----IVEMKEPKLAENLPPOTP-----VCTLATITPEYTG--PEPMVS 1399
 1125 ---YYS--GMTGPHDFYNRGSGPAPGP---VVLNSLVNDVACEKTK-----Q 1165
 1400 KILFYSASLTERKYGAFHETAPDMLPLAHVALLIGAMHRLARALEVPDIDIRQVRHQ 1459
 1166 IEGDQSMLEPOYCTTIRKANIN---GRVLAQCNIDELKEMNN-----F 1207
 1460 LVGLEAVTILRIWIFLKEQATPVALRSALQIRGRDVARREMDRAKLDLGGTPVSHIS 1519
 1208 GDMHFRSTVYLEM-----RNAE---SHVVEDPRFLSSSG-PAPHEPARASHNLT 1257
 1520 GBSITSLLELVAGRRRAEYTMAGQRLADPFOOVGYVGTGEDPEEPQSFHEE 1579
 1258 PTELSQTPYTLNFSFEELNLTGLDEGAPRHSNLSWOSQTRTPSLSSNQSDDSIETS 1317
 1580 EEVAVS-----EIRTVVRE--RHVHSENGPIVEERITTYEDDVAVNE 1624
 1318 KLUDY-----QAEYRAYREYIAMSQOLEGSGSTIISGRSPHSHTYIGOSSSGS 1370
 1625 EYDVKIVPLNEEBOEKMDRVRVENENFBOE---TSKESTFCQCTHREKDDGGS 1679
 1371 IHSNLEOERKQSEKPP---DDGRKSFLKRGDVIDYSSSGYSTNDASP-LDPIEED 1424
 1680 LKTTM-----KDSHVQIFPDGEGTS-----ANETGLSSDDADIMPTTKED 1722

RESULT 6
 US-09-577-034-33
 ; Sequence 33, Application US/09557034
 ; Patent No. 6365569
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang, Liang
 ; Blehm, E. Scott
 ; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
 ; PROTEINS, NUCLEIC ACID MOLECULES, AND
 ; USES THEREOF
 ; NUMBER OF SEQUENCES: 171
 ; CORESPONDENCE ADDRESS:
 ; ADDRESSEE: Carol Talkington Verser, Ph.D.
 ; Heska Corporation
 ; STREET: 1825 Sharp Point Drive
 ; CITY: Fort Collins
 ; STATE: Colorado
 ; COUNTRY: USA
 ; ZIP: 80525
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: Windows 95
 ; SOFTWARE: Wordperfect for Windows, Version 7.0
 ; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/557, 034
FILING DATE: 21-Apr-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/065, 474
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Verser, Carol Talkington
REGISTRATION NUMBER: 37,459
REFERENCE/DOCKET NUMBER: HW-5-C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 970/493-7272
TELEFAX: 970/484-9505
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 1745 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 33:
US-09-557-034-33

Query Match 5.4%; Score 480.5; DB 4; Length 1745;
Best Local Similarity 20.3%; Pred. No. 5.6e-32;
Matches 353; Conservative 234; Mismatches 602; Indels 549; Gaps 67;

QY 14 EENIPALKALLECKVDYDERNEGGOTPLMAEOGMEYKELIKGANCLEDLDN-- 71
DB 207 KDDDTKATLLQNEHNSDVTSGSFTPLHAAHYENVAOALLEKGAANNQARHNIS 266
QY 72 -----W----- -TALISAKEGHYIYEELIKCG----- 94
DB 267 PLVATKGRITNMVSLILAHGAVIDCRTLTLPLHCASSRGHQVVDLLKCAPISAK 326
QY 95 -----VLEH----- -RDMGWTALMW 110
DB 327 TKNGIPLHMAAOVDYDVTPLPLHAAHCGHVRVAKLLDRNADPNARALNGFTPLHI 386
QY 111 ACYGRDGVYELLISHGANPSVTGLQSVYPIIMAGRGHADIYHLLQNCARKNCSDKY 170
DB 387 ACKNRKIKIVLELLKYHAILEAT--TESGLPLHVAEFGALNIVYLLQOGANADVAVR 445
QY 171 GTPLPMAARKGHECEVHLLAMGADVDOEGANSMTALIVAKGYTQSVKEILKRPNV 230
DB 446 GETPLHIAARANDIVAVYVRNGAOVDAARELOTPHLIASRIGNDIVILLQANASP 505
QY 231 NLTDKGMATLMAKSEGTETIYODLDAGTYVINIPDSGPTVLIGAVRGHVEIVALL 290
DB 506 NAATRDLYTPLHIAKEGQEEVAAILMDHGTDKTLTKKFTPLHAKYGNLPAKSL 565
QY 291 OKYADIDIRGODNTALWAVEKGMATVNRDILQCNPTLEICTDGEFTPLIKATKRMNE 350
DB 566 EKGTPVDIEGKNQVPLHVAAHYNDKVALLLLENGASAAHAAKGTPLHIAAKKNOMD 625
QY 351 VVELLDGAKVSAVDKKGDTPLHIAIGRSRKIAELLRN----- 391
DB 626 IASTLLHKKANNAASKAGFTPLHIAQEGHREMAALLIENGAKVGAQARNGLPMLCA 685
QY 332 -----PKGRLLYRPNKAGETPNNDSCSHOKSILQIPARILSTPEIDGMDLG 440
DB 686 QEDRVSAVEELVKEAALDPKTKAGYTPLHVAHCH-----EG-----QINVR 727
QY 441 YDLSSALADILSEPTMQPICVGLYAQMSGKSLKLEDEMKTFAQOIEPLPFSW 500
DB 728 FLIEHGAKVSVITRSTYR-----LHQAQOCHNSVVR----- 760
QY 501 LTVFLTLGLCGILGLFAFYVHPNLGIASVLSFALLIIFPLVIYFGRRGESMAMAV 560
DB 761 -----YLIEHGASPNVHTSTGCTPLS----- 781
QY 561 LSTRARHIGYLELLKLMEFVNPPLPBOGTAKLVRF-----LFTDYNRLSSV 609
DB 782 ----IAERLGYVSAEALKTITETVITETTVTEERYKPNPEAMNETMFS-----SED 833

QY 610 GGETSLAMIAATLSDACERERGF--ATRLRFVEKTEPTDQCK----- 650
DB 834 EEDNOITANAHADPSLSLTKGLDSTGVHILHATETPLSRSPVEGTDGDDALIRKA 893
QY 651 -----KMKTCCLPSFVILFI-----ICGIIIGTLLAIF 681
DB 894 QHEPITTMADPSLSDASLPDNYTIMRTMQPSFLISFVWDARGAMRCRSHGVAILIP 953
QY 682 RVDPRHLVNAVLISTASVGLAFVLCNRTMQVDSILN-----SQKRLLNNAASKLHL 737
DB 954 RKAQPTRV-----TCR--YLGRDKLAHPPLSEGBAL--ASRILEM 991
QY 738 KSEG--FMKVLCEVEYELAMAKTIDSFONOTRIYVIL--DG-----LDACEODKV 785
DB 992 APBGAKFLEPVILEVPHFASL-----RGREYVILSRDGGQHKMKHEQLEATE-DAV 1042
QY 786 LQMLD-----TVRLESKSPFLTAIPASDPHIIKAINQNLNSVLD 827
DB 1043 QEVLNMFDAEELSOLDLHTSRITRLINDPDM--YFA-----VTVRROEVHCVGEG 1095
QY 828 NINGHYMNIV-HLPVFLNSRGLSMARKELVTSATNGDVPDSTGTGIEDADRVSONS 886
DB 1096 GV-----ILSSVPHVOAIFPDGSLT--KTIVSOAOPVPOEITRLH----- 1137
QY 887 IGEKTKGSKTALNRDIT--RRQMQRTITRQMSFDLTK--LVTEDWFSDISPQWR 941
DB 1138 -----GNRVASPLVYVPRRRKFKHTICLIPQSSNKGMLTOYSGQGPQPTLR 1190
QY 942 RLL-----NIVSYGRLRANOISN-----W-----DRLSWT-N 971
DB 1191 LILCSKTGSSPAAWEDITGTLQTLTGEVSEFTTVSARFWMDCOTPRDAARMAOEYVN 1250
QY 972 LTEQWYRYSWMLVLYEETEGIDQMTL-----KTI--YERISKNIPTTKDEPL- 1019
DB 1251 EATVPRYMAKFL-FARKTPPAGQRLRCMDDRDKLLEOERIE-IANSKOEVLVS 1308
QY 1020 -----LEIDGIRNFVFLSSRTPVLVARD--VKVLEPCTVN--LDPLR-----ETIA 1064
DB 1309 GRHQFLEFSGNL-----LPIKSGDQLSLYFLPQENRLAFMVKIRHTDNETAA 1358
QY 1065 DVRAARQISITGLATPPLPLHGGPRPABGYSQPSVCSSTSFNPRFPGVGVSPHSS 1124
DB 1359 DGR-----IVMKBEKILAEMLPQTP-----VCLATITLPEYTG--PEPMVS 1399
QY 1125 ----YYS--GMTGQHPRYNRSGPAPCP--VVLNSLNDVACEKIL-----Q 1165
DB 1400 KILFYSASLITEKYVGFHETAPRDNPLAHVALLIGADHRLARALEVDDIDIRQVRH 1459
QY 1166 IEGLDQSMRPOYCTTIKKANIN-----GRVLAQCNIDELKEMMN-----F 1207
DB 1460 LVGLEAVTILRIWFLKKEQATPVALSALQIRGRDVAVEREMRAKLDLGEPTVSHIS 1519
QY 1208 GDMHLFSTYLEM-----RNAE-----SHVVEPDRFLESSESSG-PAPHGEPARASHNEL 1257
DB 1520 GPSITLSTLLEAAGDRRRAHEVYMAQOIRAOEFPQOVYKGTTPDPEPKQSHHEE 1579
QY 1258 PHTLSSQPTYLINFSEELNTLGLDGCAPRHSNLWSQOTRRTPSLISLNSDSSIEIS 1317
DB 1580 EVAVS-----EIRTVRTE--RHVDSBNGPIVERERTITTYEDVAVNEE 1624
QY 1318 KLNDK-----QAEYRDAYREYIAQMSLEGPGSTISGSSPHSTTYMGOSSSGS 1370
DB 1625 EIVDKIVPLNEEEDKMDRMVREKMNFEQOE-----TSKEGFGCQTHEKEDDGS 1679
QY 1371 IHSNLEDEKQDSRKP--DDGRKSPFLMRKGVYIDVSSGVSSTNDASP-LDPTIED 1424
DB 1680 LKTM-----KDSHVROIFFDGETS-----ANETGLSSGADMTIMPTTKED 1722

RESULT 7
US-09-082-059-2
; Sequence 2, Application US/09082059A

Patent No. 6225086
GENERAL INFORMATION:
APPLICANT: MORROW, Jon S.
APPLICANT: Devarejan, Prasad
TITLE OF INVENTION: No. 6225086el Ankyrin Proteins and a Method for their Identificat
FILE REFERENCE: 44574-5002-US
CURRENT APPLICATION NUMBER: US/09/082,059A
CURRENT FILING DATE: 1998-05-21
EARLIER APPLICATION NUMBER: 60/047356
EARLIER FILING DATE: 1997-05-21
NUMBER OF SEQ ID NOS: 19
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 2
LENGTH: 1088
TYPE: PRT
ORGANISM: Homo sapiens
US-09-082-059-2

Query Match 4.9%; Score 438; DB 4; Length 1088;
Best Local Similarity 30.5%; Pred. No. 1.2e-28;
Matches 120; Conservative 73; Mismatches 196; Indels 4; Gaps 3;

14 EEENIPALKALECKDVERNECCGTPLMTAAEGNLEIYKELIKGANCNLEDLDMNT 73
30 KNRKLEWELLKKGASIQAVTESGLTPIHVAAGHVNIVSOLMHGASDNTNVRGET 89
74 ALISAKRGHVIYVELLKCQVNEHRDMGWTALMWCYKGRDVEYELLSHGANP-SV 132
90 ALHMAARSGQAEVNVYLVQDGAOYEAkakKDDQTPHISARIGKADIVQQLQOGASPM 149
133 TGLQSVYPIIWAAGRHADIVHLLQNGAKVNSDKYGTTPPLVMAARKHLECVKHLA 192
150 TTSGYT--PLHLSAEGHEDVAAPFLDHGASLITTKGFYPLHVAARKYGLFEVANI 207
193 MGADVDGANSMTALIVAKGYTOSVKEILKRNPNVNLTKDGTALMASKEGHET 252
208 KASPDAGKSGPLPLHVAANYDMQKVALLLDQASHPAAKAGYTPHIAAKNOMDI 267
253 VOGLDAGTYNIDRSGDTVLIGAVRGHVEIYRALLQYADIDIRGQDKTALYAVE 312
268 ATTLETGADANAVTRQIGASVHLAAGEGHVDVSLIGRANVNLNKSGLTFLHLAAQ 327
313 KGNATWBDILQCNPDREICTKDETPLIKATKMRNIEVEVELLDKAGKSAVDKGDTP 372
328 EDVNVAVELVNOGANDADQTKMGYTPHVGCHGNKIYFELLOHSAKNAKTKNGYTP 387
373 LH-TAIRGRKRLAELLRNPKDGRLLYRPKA 404
388 LHOAAOOGHTHIINVLLQNNASPMELTVNGNTA 420

US-08-810-712-10
Sequence 10, Application US/08810712G
Patent No. 6160106
GENERAL INFORMATION:
APPLICANT: Yeda Research and Development Co. LTD
TITLE OF INVENTION: Tumor Suppressor Genes, Proteins Encoded Thereby and
FILE REFERENCE: sequence list
CURRENT APPLICATION NUMBER: US/08/810,712G
CURRENT FILING DATE: 1997-03-03
EARLIER APPLICATION NUMBER: PCT/US94/11598
EARLIER FILING DATE: 1994-10-12
NUMBER OF SEQ ID NOS: 31
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 10
LENGTH: 1423
TYPE: PRT
ORGANISM: Homo sapiens
US-08-810-712-10

Query Match

4.6%; Score 409.5; DB 4; Length 1423;

Best Local Similarity 22.0%; Pred. No. 6.4e-26;
Matches 225; Conservative 149; Mismatches 338; Indels 309; Gaps 36;

5 ISGVINVEEENIPALKALECK--DVERNECCGTPLMTAAEGNLEIYKELIKGNA 62
336 VMKIIHAIINDVNPGLQHLIGLSINVDQPNKHGTPPLHIAAGCGNIIQLIKRGS 395
63 NCNIEDLDMNTALISASKRGHVIYVELLKCQVNEHRDMGWTALMWCYKGRDVEY 122
396 RIDVDGGSNAVYMARHGHVDTLKPLSENKCPDLVDKDSGMALHVAARYGHADY 455
123 LSHGANPSYTGLOYSYVPIIWAAGRHADIVHLLQNGAKVNSDKYGTTPPLVMAARK 182
456 TCASAOIPISRIKEETPLHCAAHGYSVAVALECGCNVIRKREGETPLTTSAG 515
183 HLECVKHLAMGADVDOEGANSMTALIVAKGYTOSVKEILKRNPNVNLTKDGTALM 242
516 YHDIVECLAHRGAD-----LNACKDGHIALH 542
243 IASREGHTEIVODLDAGTYVNIPIDSGDTVLIGAVRGHVEIYRALLQYADIDIRG 302
543 LAVRRCOMEVYIKTLISQCFVDYQDRHGMTPLHVACKDGNMPTVALCE----- 591
303 NKTALYAVAKGNATWBDILQCNPDREICTKDETPLIKATKMRNIEVEVELLDKGA 362
592 -----ANCNLD--ISNKTGRPLHLAANGILLDVAVRTICLMGASV 629
363 SAVDKGDPPLHIAIRGRSKRLAELLRNPKDGRLLYRPNAKGEPPYNIDCSHOKSIL 422
630 EALTIDGKTAEDLASEOHEHVAAGLLARLKD-----THRGLEIQ 670
423 IFGARHLSPEITDGMIDGLDYSSALADLSEPTMOPICVGLYAMWGSGKFLKLE 482
671 -----LRPIQ-----NLQPRKLKLFHSGSGSKTTLVESLK- 701
483 EMKTFAGQGLEPLFQSWLIVETLLLCGLGLLFAFVHPNLG----- 526
702 -----CGLNSFR--RRPRLSSTNSRPPSPPLASKP 733
527 -IAVSLFLALYIFIVYIFGGRGSESMNNAWISTLAHIGYLELLKLMVNPPE 585
734 TVSVSINNL-----YPCENVSVRSSMPEGLTN--GMLE-----VVAATG 775
586 LP-----EQTAL-----PVRLFTDYRRLSSVGETSLAMEI 619
776 HPHCSADQSTFAIDIQNAYLNGVDGFSVWEFSNPVYFCYD--PAANDPISIHVV 832
620 ATLSDACEREG--FLATRLFRVETEDT--GCKKKKTKCCLPSFVILEITIGCITS 673
833 FSLDEPEYETQLNPVIFWLSFLKSLVPEEPIAFGGKLN-----NPLQVVL----- 877
674 GITLLAIFRVDPRKHLTVNAVLSIASVYGLAFVLCRPMQVLDLSOKRRLHN--A 730
878 -----VATHADIMNVPRPAGGEGYDKT-----SLKEINRRGNDLHI 917
731 ASKLHLKASEGF---MKVLKCEVELAMAKTIDSFQNTQTRLVVIIDGACEDQV 786
918 SKLFLVLDAGASGSKDMKVLRNHQ--EIRSOIYVCPPTMHL-----CE--KII 963
787 QMLDVRVLFESKGPPIAL--FASDPHIIKAIQNLNSVLDPSNINHDIYKRNIVHLPV 844
964 STLPFWRKINGPNQMLSDQFVYD-----VOQDLNPLASEED-----LRRTAQD-- 1007
845 INSRGLSNARKFLVYSAFNGVPCSDTGTIGEDADRRVSQNSLSEMTKIGSKTALNR-RD 903
1008 LHSITGEIN-----IMSETIVQVLLD-----PRMLCTNVLLKLSVETPRALHHRG 1055
904 TYRRROKQRTITRQNSFDLTKLVTEWFS--DISPQTRRLNLIVSYTGRLIRANQISFN 962
1056 RYVEDIQLRVPDSQVEELQIILDAMDICARDLSSGTM-----VDVPA-LIKTDNLRHS 1108
963 W 963

Db 1109 W 1109

RESULT 9

US-09-196-387-2

Sequence 2, Application US/09196387

Patent No. 6277613

GENERAL INFORMATION:

APPLICANT: de Lange, Titia

APPLICANT: Smith, Susan

TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS

TITLE OF INVENTION: OF USE THEREOF

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESS: Klauder & Jackson

STREET: 411 Hackensack Avenue, 4th Floor

CITY: Hackensack

STATE: New Jersey

COUNTRY: USA

ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/196,387

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/095,225

FILING DATE: June 10, 1998

ATTORNEY/AGENT INFORMATION:

NAME: Jackson Esq., David A.

REGISTRATION NUMBER: 26,742

REFERENCE/DOCKET NUMBER: 600-1-230 CIP1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201-487-5800

TELEFAX: 201-343-1684

TELEX: 133521

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1327 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

US-09-196-387-2

Query Match

Best Local Similarity 22.3%; Pred No. 4.4e-23;

Matches 100; Conservative 89; Mismatches 220; Indels 247; Gaps 17;

Db 14 EENIPALKEKCKVDNERGCGOTPLMAAEGNEIYELIKNGANCNLEDDMNT 73

Db 348 EKKIALLPLPLVNCNCHSDGRK---STPLHLAAGYNNRYIVOLLHGHGADVAAKKGGLV 404

Db 74 ALISAKSEGHVHIVVELLKCGVNLHRDGMGTALMMACYKRTIVVELLSHGANSPT- 132

Db 405 PLHMACSYGHEVTELLKNGACVNAADLMOTPLHMAASKRVEVCSLLSHGADPTLV 464

Db 133 -----T 133

Db 465 NCHGSAVDMAPTPELRELTVEFGHSLDQAAREADLAKVKTALLETINFKOPSHET 524

Db 134 GLQYSV-----YPTTAAARGHADIIVHLLONGAKV 164

Db 525 ALHCAVASLHPRKKQYTELLRKGANVNEKNRDFMTPLHVAABRAHNDVMEVLHKGAKM 584

Db 165 NCSDKYGTPPLVMAAKGHLCEYKHLAMGADVDDGAGNSMTA----- 207

Db 585 NALDTLGGTALHRAALAGHLQYCRLLSLYSGPSITISLOGFTTAOMGNBAVVOIISSTP 644

QY 208 -----LIYAVKGYTQSVKEILKRNPNVNLIDKDS--NTALMASKEGHEIYDVL 256

Db 645 IRTSDVDYRLLEKSKAGDDEYKQLCS-SQNNVNCRLBGRHSTPLHFAAGYKRVSVLEL 703

QY 257 LDAGTYVNIPIRSGDFTVLIGAVRGHVEIVRALLQKYADIDIRGDNKTALYWAYEK-- 314

Db 704 LHHGADVAAKKGGLVPLHMACSYGHEVEALIVRHGASVNAADLMKFTPLHMAAKGY 763

QY 315 -----NATVVRILQ-----C 325

Db 764 EICKLLKGGADPTKKNRDNPTLVKSGDIDILGDAALLDAKGLARVQKLC 823

QY 326 NPDEICTKD---GETPLKATMRNIEVELLDKGAVSADVKKGDTPLH----- 374

Db 824 TPEINCRDPQGNSTPLHAAQYNNLEVAEYELHGGADVNNQDKGLPLHMAASGYHV 883

QY 375 -----TAIRGRSKLAELLRNPKDGRLLYRPNKAGETP 408

Db 884 DIAALLIKYNTCVNATDKNFATPLHBAOKGR-QLCALLHAGADPTM---KNQGGTP 939

QY 409 YN-----IDCSHOKSILT-----QIFGARHLSPTEPDGMDGLDYSSALADI 451

Db 940 LDLATADDIPLALIDAMPPLPCTCKRPOATVVSASLISPASTPS-----CLSAASIDN 994

QY 452 LSEPTMQPICVGLYAQWGSKSFLLK-----LEDEKTRPAGQIEPLFQ 497

Db 995 LTGPIAE--LAVGASNAAGDGAAGTERKEBEVAGLDMMNISQFLKSLGLEHLRDI 1048

RESULT 10

US-09-031-485-28

Sequence 28, Application US/09031485

Patent No. 5824306

GENERAL INFORMATION:

APPLICANT: Tang, Liang

APPLICANT: Blehm, E. Scott

TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN

TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND

TITLE OF INVENTION: USES THEREOF

NUMBER OF SEQUENCES: 85

CORRESPONDENCE ADDRESS:

ADDRESS: Carol Talkington Verser, Ph.D.

ADDRESS: Hesk Corporation

STREET: 1825 Sharp Point Drive

CITY: Fort Collins

STATE: Colorado

COUNTRY: USA

ZIP: 80525

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: Windows 95

SOFTWARE: WordPerfect for Windows, Version 7.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/031,485

FILING DATE:

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/847,429

FILING DATE: 24-APR-1997

ATTORNEY/AGENT INFORMATION:

NAME: Verser, Carol Talkington

REGISTRATION NUMBER: 37,459

REFERENCE/DOCKET NUMBER: HW-5

TELECOMMUNICATION INFORMATION:

TELEPHONE: 970/493-7272

TELEFAX: 970/484-9505

INFORMATION FOR SEQ ID NO: 28:

SEQUENCE CHARACTERISTICS:

LENGTH: 348 amino acids

TYPE: amino acid

TOPOLOGY: linear

US-09-065-474-28

TOPOLOG: LINEAR

SEQUENCE CHARACTERISTICS:
LENGTH: 949 amino acids

TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-196-387-10

Query Match 4.2%; Score 370.5; DB 4; Length 949;

Best Local Similarity 24.1%; Pred. No. 7.7e-23;
 Matches 139; Conservative 62; Mismatches 174; Indels 201; Gaps 11;

14 EENIPALKEKEDKDERNECQGPLMAEEOGNEIYKELKNGANCNLEDLDWMT 73
 348 EKKMALPLPLVNVNCHASGRK---STPLHLAAGNRRRIYOLLQHGADVHAKDGGLV 404
 74 AIISSKSGHVIIVELLKCGVNLHRDMGCGTALMAQYGRDVBELLSHGANSV- 132
 405 PLHNACSYGHEVETELLKHGACVNAAMDLOFTPLHEAASKNRVVCSTLLSHGADPTLV 464
 133 -----T 133
 465 NCHGSAVDMAPELRERLYEFGSHSLQAAAREADLAKYKTLALEIINEKOPQSHET 524
 134 GLQYV-----YPIIWAAGRHADIYHLLONGAKV 164
 525 ALHCAVASLHPKRGQVTELLRKGANVEKNKDFPTPLHVAERAHNDVMEYLKHKAKM 584
 165 NCSKRYGTPPLVMAARKGHECVKHLAMGADVDOEGANSMTA----- 207
 585 NALDTLGGTALHRAALAGHLQTCRLLSYGSDPSIISLOGFTAAQMGNEAVQQLSESTP 644
 208 -----LIIAIVGGYTSYKELKRNPNVNLTDKCG--NTALMISKSGHEIYQDL 256
 645 IITSVDYRLLEASAGLETVKQLCS--SQVNCDELGRHSTPLHFAAGYNRVSVETL 703
 257 LDAGTYVNIIDRSQDVLIGAVRGHVEIVRALLQKADIDIRGQDNKTALYWAVERG-- 314
 704 LHHGADVNAKDKGLVPLHNACSYGHEVAELLYVHNGASVNAVADLMKTPPLHEAAKKY 763
 315 -----NATWVDLIQ-----C 325
 764 EICKLLHNGADPTKKNBDGNTPLDLVKEGDTIDQLKGDALDAALDAKKGLARVOKLC 823
 326 NEDTEICTKD--GETPLIKATKRNIEVVELLDKAGKYSADVKKGDPPLHIAIR----- 378
 824 TPENINCNDQGRNSTPLHLAAGYNNLEVAEYLLLEHGDVNAQDKGLIPLHNAASVGGC 883
 379 -GRSRKLAELLRNPKDGRLLYRPNAKGETRYNIDC 413
 884 IARVOKLC-----TPENINC 898

RESULT 15

US-09-350-982C-5
 Sequence 5, Application US/09350982C
 Patent No. 6455290

GENERAL INFORMATION:

APPLICANT: Berthelsen, Jens
 APPLICANT: Toma, Salvatore
 APPLICANT: Isacchi, Antonella

TITLE OF INVENTION: Tankyrase Homolog Protein (TNP), Nucleic Acids, and Methods Relat

FILE REFERENCE: PHRM-0043

CURRENT APPLICATION NUMBER: US/09/350,982C

NUMBER OF SEQ ID NOS: 10

SOFTWARE: PatentIn version 3.0

SEQ ID NO 5
 LENGTH: 1166

TYPE: PRT
 ORGANISM: Artificial

NAME/KEY: misc_feature
 OTHER INFORMATION: Xaa is any amino acid
 NAME/KEY: misc_feature

LOCATION: (1102)..(1102)
 OTHER INFORMATION: n is any nucleic acid
 NAME/KEY: misc_feature
 LOCATION: (2650)..(2650)
 OTHER INFORMATION: n is any nucleic acid
 US-09-350-982C-5

Query Match 4.1%; Score 363; DB 4; Length 1166;

Best Local Similarity 22.9%; Pred. No. 5.2e-22;
 Matches 141; Conservative 65; Mismatches 181; Indels 228; Gaps 12;

19 PALKALLEKCKD-----VDERNECG--TPMIAEEOGNEIYKELIKNGA 62
 23 PAAREFEACRNDVERKRLVPEKYNSTHDTGRKSTPLHLAAGGRDVEYELLQNGA 82
 63 NCNLEDNDNTALISAKSGHVIIVELLKCGVNLHRDMGCGTALMAQYGRDVBEL 122
 83 NVQARDDGGILPLHNACSFHAEVYVNLHNGADPNARDWNTPLHEAALIKIDVCTIV 142
 123 LLSHGANSV-----TGLQY 137
 143 LQHGAEPTIRNTDGRALDLADPSAKAVLTGEYKKDELLESARSGNEKEMALLTPLAV 202
 138 SVT-----PIIWAAGRHADIYHLLONGAKVNCSDRYGTPPLVMAARKGHE----- 185
 203 NCHASDKRSTPLHLAAGYNRKRYVOLLQHGADVHAKDKGLVPLHNACSYGHEVTEL 262
 186 -----CY----- 187
 263 LKHGACVNAAMDLOFTPLHEAASKNRVVCSTLLSYGADPTLLNCHNSAIDLAPTPOL 322
 188 -----KH----- 189
 323 KERLAIEFGKSHSLQAAAREADYTRIKHLSLEWVNFKHQTHETAXHCAASPYPRKQI 382
 190 ---LIMGADVDOEGANSMTALVYAKGGTQSYKELKRNPNVNLTDKDGNTALMIASK 246
 383 CELLRKGAXINKEKTEFTPLHVASERAHNDVEVYVNHAKVNALMDLIGOTSLHRAV 442
 247 EGH-----TEIVODLDAGTYVNIIDRSQDVLIGAV 278
 443 CGHLQTCRLLSYGCDPNIIISLOGFTALQMGNNVOQLQEG--ISLGSNDROLLENA 500
 279 RGHVEIVRAL--LQRYADIDIRGQDNKTALYWAVERGNATYVRDILQCNPTETICTKG 336
 501 KAGDVEYVKLCITVQSYNCRDIEGROS--TPLHFAAGYNRVSVETYLQHGADVHAKDKX 559
 337 ETPLIKATKRNIEVVELLDKAGKYSADVKKGDPPLHIAINGRSKLAELLRNPKGR 396
 560 LVPLHNACSYGHEVAELLYKGAVVNAVADLMKFTPLHEAALKGKEYICKLLQHGADPT 619
 397 LLYRPNAKGETPYNI 411
 620 ---KKNRDGTPLDL 631

Search completed: July 1, 2003, 14:41:04
 Job time : 39.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2003, 14:34:04 ; Search time 16 Seconds
(Without alignments)
4445.744 Million cell updates/sec

Title: US-10-021-571-2

Sequence: 1 MSVLISQSVNYVEENIPR.....ELHAASSESTGCEERESITL 1715

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	556	6.3	1880	ANK1_HUMAN	P16157 homo sapien
2	550	6.2	3924	ANK2_HUMAN	Q01484 homo sapien
3	527.5	6.0	1862	ANK1_MOUSE	Q02357 mus musculu
4	526	5.9	4377	ANK3_HUMAN	Q12955 homo sapien
5	468	5.3	768	VR23_HUMAN	Q91177 homo sapien
6	431.5	4.9	1059	VR379_HUMAN	Q15064 homo sapien
7	414.5	4.7	1431	DAPK_HUMAN	P53353 homo sapien
8	384.5	4.5	1327	ANK1_HUMAN	Q95271 homo sapien
9	384.5	4.3	1166	ANK2_HUMAN	Q96272 homo sapien
10	380.5	4.3	1166	ANK2_HUMAN	Q96272 homo sapien
11	378.5	4.3	323	ANKH_CHRYT	Q06527 chromatium
12	360	4.1	832	ANK3_HUMAN	P57078 homo sapien
13	357.5	4.0	741	RNSA_HUMAN	Q05823 homo sapien
14	356	4.0	735	RNSA_MOUSE	Q05921 mus musculu
15	349.5	3.9	692	ANK6_HUMAN	Q05294 homo sapien
16	335	3.8	1401	LATA_LATWA	P26331 latrodectus
17	331	3.7	525	ASB3_MOUSE	Q94772 mus musculu
18	329	3.7	583	ASB5_MOUSE	Q94772 mus musculu
19	328.5	3.7	518	ASB3_HUMAN	Q95755 homo sapien
20	328.5	3.7	747	V222_FOWPV	Q91513 fowlpox vir
21	310.5	3.5	542	V155_FOWPV	Q91517 fowlpox vir
22	309	3.5	668	V244_FOWPV	Q91426 fowlpox vir
23	308	3.5	596	V024_FOWPV	Q91517 fowlpox vir
24	307	3.5	436	V245_FOWPV	Q91425 fowlpox vir
25	288	3.3	525	V228_FOWPV	Q91507 fowlpox vir
26	288	3.3	1083	V112_YEAST	P40480 saccharomyc
27	271	3.1	592	V246_FOWPV	Q91424 fowlpox vir
28	270	3.0	603	V162_FOWPV	Q91569 fowlpox vir
29	265.5	3.0	434	AS15_HUMAN	Q84841 homo sapien
30	264.5	3.0	656	FEHL_CARTEL	P17221 caenorhabdi
31	260	2.9	433	AS14_MOUSE	Q84847 mus musculu
32	254.5	2.9	776	ANK5_HUMAN	Q94802 homo sapien
33	251	2.8	429	AS10_HUMAN	Q84813 homo sapien

34	245.5	2.8	578	1	V022_FOWPV	Q91517 fowlpox vir
35	245	2.8	278	1	AS13_HUMAN	Q84843 homo sapien
36	242	2.7	434	1	V023_FOWPV	Q91517 fowlpox vir
37	238.5	2.7	333	1	ANK2_HUMAN	Q94771 homo sapien
38	237.5	2.7	1066	1	NDU2_MOUSE	Q94771 mus musculu
39	236.5	2.7	328	1	ANK2_MOUSE	Q94771 mus musculu
40	235.5	2.7	2703	1	NORC_DROME	P07207 drosophila
41	235	2.7	231	1	PSDA_RAT	Q94771 ratius norv
42	234	2.6	642	1	YAZA_SCHPO	Q94771 schizosacch
43	233	2.6	437	1	V014_FOWPV	Q91517 fowlpox vir
44	230.5	2.6	461	1	V218_FOWPV	Q91517 fowlpox vir
45	230	2.6	226	1	PSDA_HUMAN	Q75832 homo sapien

ALIGNMENTS

RESULT 1

ID ANK1_HUMAN STANDARD; PRT: 1880 AA.

AC P16157; 01-APR-1990 (Rel. 14, Created)

DT 01-APR-1990 (Rel. 14, Last sequence update)

DE 15-JUN-2002 (Rel. 41, Last annotation update)

DE Ankyrin 1 (Erythrocyte ankyrin) (Ankyrin R).

GN ANK1 OR ANK.

OS Homo sapiens (human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

NCBI_Taxid=9606;

RP [1]

RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND VARIANTS.

RC TISSUE=Hematopoietic;

RA MEDLINE=90158830; PubMed=2137557;

RA Lux S.E., John K.M., Bennett V.;

RT "Analysis of cDNA for human erythrocyte ankyrin indicates a repeated

RT structure with homology to tissue-differentiation and cell-cycle

RT control proteins.";

RL Nature 344:36-42(1990).

RN [2]

RN SEQUENCE FROM N.A.

RX MEDLINE=9015370; PubMed=169849;

RA Lambert S., Yu H., Pichal J.T., Lawler J., Ruff P., Speicher D.,

RA Cheung M.C., Kan Y.W., Palek J.;

RT "cDNA sequence for human erythrocyte ankyrin.";

RT Proc. Natl. Acad. Sci. U.S.A. 87:1730-1734(1990).

RN [3]

RN VARIANT HS ILE-462.

RX MEDLINE=96225450; PubMed=8640229;

RA Eder S.W., Gonzalez J.M., Lux M.L., Scarpa A.L., Tse W.T.,

RA Dornwell M., Herbers J., Kugler W., Oezcan R., Pekrun A.,

RA Gallacher P.G., Schroeter W., Forget B.G., Lux S.E.;

RT "Ankyrin-1 mutations are a major cause of dominant and recessive

RT hereditary spherocytosis.";

RL Nat. Genet. 13:214-218(1996).

RL -1- FUNCTION: ATTACH INTEGRAL MEMBRANE PROTEINS TO CYTOSKELETAL

RL ELEMENTS. BIND TO THE ERYTHROCYTE MEMBRANE PROTEIN GP85, AND TO THE

RL NA-K ATPASE. TO THE LYMPHOCYTE MEMBRANE PROTEIN GP85, AND TO THE

RL CYTOSKELETAL PROTEINS FODRIN, TUBULIN, VIMENTIN AND DESMIN.

RL ERYTHROCYTE ANKYRINS ALSO LINK SPECTRIN (BETA CHAIN) TO THE

RL CYTOSKELETAL DOMAIN OF THE ERYTHROCYTES ANION EXCHANGE PROTEIN;

RL THEY RETAIN MOST OR ALL OF THESE BINDING FUNCTIONS.

RL -1- SUBCELLULAR LOCATION: CYTOPLASMIC SURFACE OF ERYTHROCYTIC

RL PLASMA MEMBRANE.

RL -1- ALTERNATIVE PRODUCTS: At least 3 isoforms; 1/2.1 (shown here),

RL 2/2.2 and 3. are produced by alternative splicing.

RL -1- P.M.: REGULATED BY PHOSPHORYLATION.

RL -1- P.M.: ACYLATED BY PALMITIC ACID GROUP(S).

RL -1- DISEASE: Defects in ANK1 are the cause of dominant and recessive

RL hereditary spherocytosis (HS).

RL -1- SIMILARITY: CONTAINS 23 ANK REPEATS.

RL -1- SIMILARITY: CONTAINS 1 DEATH DOMAIN.

1

[illegible]

```

QY 613 TSLAEMIANLSDACEREFGLATRLR-----VPRTEE-SQKKKKKKTKCCLPSPV 662
DB 840 KELLDEVPKLDQVE-----SPAIPRICAMPETVIVSEEDQSKKEYDESLIPS-- 891
QY 663 IFLFIYGLIAGITLLAIFRVDPKHLTVNAILISIASVGLAFVLCRTMVOYLISLNS 722
DB 892 -----SPATETSDNI-SPVAPSHVHTGLVAF-----WVDARGGS 924
QY 723 QRRRLHSAASKLHLKSEGFMYKCEVELAMAMAKTIDTFQONQRLVIYIDGLDACIQ 782
DB 925 MRSRHN-----GLRVVIP-----PRTCAAPRTICRLV----- 953
QY 783 DKVLQMLDYRVLFSGKPIAIFASDPHIIKAINONLNSVLDNSNINGDMWRN-IYHL 841
DB 954 --KPKLSTPPPL-----AEEGLASRIITALPGTGAQPLSPYIVEL 992
QY 842 PVFL-NSRGLSNARKFLVTSATNGDITCSPTGTQEDTDORVSONSIGEWTKLSKTAAL 900
DB 993 PHFASHGRG---DRELVLVLSKENGSV-----WHEHRSRGESYLQDILN-GMDEBELG 1040
QY 901 RRDYRRRQOMRTTROMSFDLTKLVT---EDWFSDISQIMRLINIVSYTGILLRAN 957
DB 1041 SLELEKRRVCRIIT--TDFPLFYVIMSRLCOO-YDIIGPEG-----GSLKSKLVPYV 1090
QY 958 QIFNMORLASWINTLEQMPYRTSWILYLEETEGLDQDQTLK-----TMYER 1005
DB 1091 QATPEPNAVTKRKALQ-----AOPVDELVTGLGNQATFSPITVEPR 1136
QY 1006 ISK-----NIPTKDEPLEIDIDIRNEVF-----LSSRTPLVA 1042
DB 1137 RRRKRRIGIRIPLPWTQNPDSRGSDTSLRLCSVIGTDQAMQEDITGTKLVA 1196
QY 1043 RDKVTEPLPCVINDPKREIITADYRAREQINIGLAPPLPHGEPFPPSGYSOPASY 1102
DB 1197 NECANF---TTNV--SARFMLSQCPRTAEAVNATLTKEL-----TAVPYAKF 1241
QY 1103 CSSASFNGPPGCVVSPQPHSSVYSGLSGQHPPYNAAPATGSSLLSSMVDVYCEK 1162
DB 1242 VIFAKMDPRBG-----RLRCYCMDDKDYDKT 1268
QY 1163 LRQ-----IEG-----LDONMPOYCTTIKKAN-----INGRVL 1191
DB 1269 LEQENHVEVARSDIDIVLEGMSLFAELSGNLV---YKKAQOQSFHFGSREHRLA 1323
QY 1192 SQCNIDELKE-----MANFGD-WHLFRSVMLEKMSVESQVPEPR----- 1233
DB 1324 MPVAVRDSRPPGSLFLRKAMKYEOTOHLCHLNTMPPCAGSGARRRRTPPLAL 1383
QY 1234 ---FLNENSSAPVPHGESAR---SSHTEPLTESSQTPYTLNFSPELNTGLD-- 1283
DB 1384 RYSLIESTPGSLGTOAEKMAVISEHGLSMAELARE---LOFVEDINIRVENP 1439
QY 1284 ---EGAPRHSNLTWQOTRTPSLSSLNQSDSISEIKLTDQVA-----BYRD 1329
DB 1440 NSLLEQSVALLNL-WVIREGONAMENLYTALQSIDGELYNMLEGGROSRLNKPRRH 1498
QY 1330 AYREYIAQMSQLEGGTSSITSGRSSPHSTYYIGOSSSGGSHSTLEQERKEGELKOED 1389
DB 1499 TDRYSLSPSQMNG--YSLQDELISP-----ASIGCALSSPLRADQ----- 1538
QY 1390 GRKSLMKRGVYIDYSSSGVSTNAPSLDPL---TEED--EKSDOS--GSKLLPKGS 1440
DB 1539 -----YWNEVAVYDALPPLATHTDHTMLEMSDQVMSAGLTPSLVT 1578
QY 1441 SERPSIFQTDLTKGGLRYQKLSDEDESGTGHVQJLTPHCSKMIKRLKAKORECASP 1500
DB 1579 AEDSSL-----ECKAKEDSDATGH-----EMKLBGALSEPRGP 1612
QY 1501 QEHSAPERTIRAKAYLSDA---LLDKDSSDGSVNSNESSPNHSLHNEAADDSQLEKA 1557
DB 1613 ELGSLIE---LYVEDTVSDATNGLIDLE--QEEGORSSEKLPKSKRQDDATGAGQ--DSE 1666
QY 1558 NLIELEDGSHGKGGMHSHLSGLQDPIIARMSISGEK-----KSPSCSLIASPESW 1612

```

```

DB 1667 NEVSLVSHGRQQRANITHS-----PTVSQYTERSQDRLQWMDADGSTVSYLQDAAGQSW 1720

RESULT 2
ANK2_HUMAN
ID ANK2_HUMAN STANDARD: PRT: 3924 AA.
AC 001484; 001485.
DT 01-APR-1993 (Rel. 25, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ankyrin 2 (Brain ankyrin) (Ankyrin B) (Ankyrin, nonerythroid).
GN ANK2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1; 2 AND 3).
RC TISSUE=Brain stem;
RX MEDLINE=91302466; PubMed=1830053;
RA Otto E., Kunimoto M., McLaughlin T., Bennett V.;
RT "Isolation and characterization of cDNAs encoding human brain
ankyrins reveal a family of alternatively spliced genes.";
RL J. Cell Biol. 114:241-253(1991).
RN [2]
RP REVISIONS.
RA Carpenter S.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Brain stem;
RX MEDLINE=94075409; PubMed=8253844;
RA Chan W., Kordell E., Bennett V.;
RT "440-kD ankyrin: structure of the major developmentally regulated
domain and selective localization in unmyelinated axons.";
RL J. Cell Biol. 123:1463-1473(1993).
RN [4]
RP SEQUENCE OF 463-495 FROM N.A.
RX MEDLINE=92009921; PubMed=1833308;
RA Tse W.T., Meuninger J.C., Yang-Feng T.L., Francke U., Sahr K.E.,
RL Lux S.E., Ward D.C., Forget B.G.;
RT "Isolation and chromosomal localization of a novel nonerythroid
ankyrin gene.";
RL Genomics 10:858-866(1991).
CC - FUNCTION: Attach integral membrane proteins to cytoskeletal
elements. Also bind to cytoskeletal proteins.
CC - ALTERNATIVE PRODUCTS: 3 ISOFORMS; 1 (SHOWN HERE), 2 AND 3; ARE
PRODUCED BY ALTERNATIVE SPLICING.
CC - TISSUE SPECIFICITY: PLASMA MEMBRANE OF NEURONS AS WELL AS GLIAL
CELLS THROUGHOUT THE BRAIN.
CC - PTM: PHOSPHORYLATED AT MULTIPLE SITES BY DIFFERENT PROTEIN KINASES
AND EACH PHOSPHORYLATION EVENT REGULATES THE PROTEIN'S STRUCTURE
AND FUNCTION (POTENTIAL).
CC - SIMILARITY: CONTAINS 1 DEATH DOMAIN.
CC - SIMILARITY: CONTAINS 23 ANK REPEATS.
CC - SIMILARITY: CONTAINS 1 DEATH DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
EMBL: X56957; CAA40278.1; -
EMBL: X56958; CAA40279.2; -
EMBL: Z26634; CAA42644.1; -
EMBL: M37123; AAA62828.1; -
PIR: S14533; S14533.
PIR: A39643; A39643.
PIR: B39643; B39643.
PIR: S14569; S14569.

```

DR HSP; P42771; IDC2.
 DR Genew; HGNC:493; ANK2.
 DR MIM; 106410; -.
 DR InterPro; IPR002110; ANK.
 DR InterPro; IPR000488; Death.
 DR InterPro; IPR000906; ZUS.
 DR Pfam; PF00023; ank; 24.
 DR Pfam; PF00531; death; 1.
 DR Pfam; PF00791; ZUS; 1.
 DR PRINTS; PR01415; ANKYRIN.
 DR SMART; SM00248; ANK; 21.
 DR SMART; SM00005; DEATH; 1.
 DR SMART; SM00218; ZUS; 1.
 DR PROSITE; PS50088; ANK_REPEAT; 20.
 DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
 DR PROSITE; PS50017; DEATH_DOMAIN; 1.
 DR Cytoskeleton; Alternative splicing; Repeat; ANK repeat;
 KW Phosphorylation.
 FT REPEAT 63 92 ANK 1.
 FT REPEAT 96 125 ANK 2.
 FT REPEAT 129 158 ANK 3.
 FT REPEAT 162 191 ANK 4.
 FT REPEAT 193 220 ANK 5.
 FT REPEAT 232 261 ANK 6.
 FT REPEAT 265 294 ANK 7.
 FT REPEAT 298 327 ANK 8.
 FT REPEAT 331 360 ANK 9.
 FT REPEAT 364 393 ANK 10.
 FT REPEAT 397 426 ANK 11.
 FT REPEAT 430 459 ANK 12.
 FT REPEAT 463 492 ANK 13.
 FT REPEAT 496 525 ANK 14.
 FT REPEAT 529 558 ANK 15.
 FT REPEAT 562 591 ANK 16.
 FT REPEAT 595 624 ANK 17.
 FT REPEAT 628 657 ANK 18.
 FT REPEAT 661 690 ANK 19.
 FT REPEAT 694 723 ANK 20.
 FT REPEAT 727 756 ANK 21.
 FT REPEAT 760 789 ANK 22.
 FT REPEAT 793 822 ANK 23.
 FT REPEAT 825 858 ANK 24.
 FT REPEAT 861 894 ANK 25.
 FT REPEAT 897 930 ANK 26.
 FT REPEAT 933 966 ANK 27.
 FT REPEAT 969 1002 ANK 28.
 FT REPEAT 1005 1038 ANK 29.
 FT REPEAT 1041 1074 ANK 30.
 FT REPEAT 1077 1110 ANK 31.
 FT REPEAT 1113 1146 ANK 32.
 FT REPEAT 1149 1182 ANK 33.
 FT REPEAT 1185 1218 ANK 34.
 FT REPEAT 1221 1258 ANK 35.
 FT REPEAT 1257 1297 ANK 36.
 FT REPEAT 1293 1333 ANK 37.
 FT REPEAT 1329 1369 ANK 38.
 FT REPEAT 1365 1405 ANK 39.
 FT REPEAT 1401 1441 ANK 40.
 FT REPEAT 1437 1477 ANK 41.
 FT REPEAT 1473 1513 ANK 42.
 FT REPEAT 1509 1549 ANK 43.
 FT REPEAT 1545 1585 ANK 44.
 FT REPEAT 1581 1621 ANK 45.
 FT REPEAT 1617 1657 ANK 46.
 FT REPEAT 1653 1693 ANK 47.
 FT REPEAT 1689 1729 ANK 48.
 FT REPEAT 1725 1765 ANK 49.
 FT REPEAT 1761 1801 ANK 50.
 FT REPEAT 1797 1837 ANK 51.
 FT REPEAT 1833 1873 ANK 52.
 FT REPEAT 1869 1909 ANK 53.
 FT REPEAT 1905 1945 ANK 54.
 FT REPEAT 1941 1981 ANK 55.
 FT REPEAT 1977 2017 ANK 56.
 FT REPEAT 2013 2053 ANK 57.
 FT REPEAT 2049 2089 ANK 58.
 FT REPEAT 2085 2125 ANK 59.
 FT REPEAT 2121 2161 ANK 60.
 FT REPEAT 2157 2197 ANK 61.
 FT REPEAT 2193 2233 ANK 62.
 FT REPEAT 2229 2269 ANK 63.
 FT REPEAT 2265 2305 ANK 64.
 FT REPEAT 2301 2341 ANK 65.
 FT REPEAT 2337 2377 ANK 66.
 FT REPEAT 2373 2413 ANK 67.
 FT REPEAT 2409 2449 ANK 68.
 FT REPEAT 2445 2485 ANK 69.
 FT REPEAT 2481 2521 ANK 70.
 FT REPEAT 2517 2557 ANK 71.
 FT REPEAT 2553 2593 ANK 72.
 FT REPEAT 2589 2629 ANK 73.
 FT REPEAT 2625 2665 ANK 74.
 FT REPEAT 2661 2701 ANK 75.
 FT REPEAT 2697 2737 ANK 76.
 FT REPEAT 2733 2773 ANK 77.
 FT REPEAT 2769 2809 ANK 78.
 FT REPEAT 2805 2845 ANK 79.
 FT REPEAT 2841 2881 ANK 80.
 FT REPEAT 2877 2917 ANK 81.
 FT REPEAT 2913 2953 ANK 82.
 FT REPEAT 2949 2989 ANK 83.
 FT REPEAT 2985 3025 ANK 84.
 FT REPEAT 3021 3061 ANK 85.
 FT REPEAT 3057 3097 ANK 86.
 FT REPEAT 3093 3133 ANK 87.
 FT REPEAT 3129 3169 ANK 88.
 FT REPEAT 3165 3205 ANK 89.
 FT REPEAT 3201 3241 ANK 90.
 FT REPEAT 3237 3277 ANK 91.
 FT REPEAT 3273 3313 ANK 92.
 FT REPEAT 3309 3349 ANK 93.
 FT REPEAT 3345 3385 ANK 94.
 FT REPEAT 3381 3421 ANK 95.
 FT REPEAT 3417 3457 ANK 96.
 FT REPEAT 3453 3493 ANK 97.
 FT REPEAT 3489 3529 ANK 98.
 FT REPEAT 3525 3565 ANK 99.
 FT REPEAT 3561 3601 ANK 100.
 FT REPEAT 3597 3637 ANK 101.
 FT REPEAT 3633 3673 ANK 102.
 FT REPEAT 3669 3709 ANK 103.
 FT REPEAT 3705 3745 ANK 104.
 FT REPEAT 3741 3781 ANK 105.
 FT REPEAT 3777 3817 ANK 106.
 FT REPEAT 3813 3853 ANK 107.
 FT REPEAT 3849 3889 ANK 108.
 FT REPEAT 3885 3925 ANK 109.
 FT REPEAT 3921 3961 ANK 110.
 FT REPEAT 3957 3997 ANK 111.
 FT REPEAT 3993 4033 ANK 112.
 FT REPEAT 4029 4069 ANK 113.
 FT REPEAT 4065 4105 ANK 114.
 FT REPEAT 4101 4141 ANK 115.
 FT REPEAT 4137 4177 ANK 116.
 FT REPEAT 4173 4213 ANK 117.
 FT REPEAT 4209 4249 ANK 118.
 FT REPEAT 4245 4285 ANK 119.
 FT REPEAT 4281 4321 ANK 120.
 FT REPEAT 4317 4357 ANK 121.
 FT REPEAT 4353 4393 ANK 122.
 FT REPEAT 4389 4429 ANK 123.
 FT REPEAT 4425 4465 ANK 124.
 FT REPEAT 4461 4501 ANK 125.
 FT REPEAT 4497 4537 ANK 126.
 FT REPEAT 4533 4573 ANK 127.
 FT REPEAT 4569 4609 ANK 128.
 FT REPEAT 4605 4645 ANK 129.
 FT REPEAT 4641 4681 ANK 130.
 FT REPEAT 4677 4717 ANK 131.
 FT REPEAT 4713 4753 ANK 132.
 FT REPEAT 4749 4789 ANK 133.
 FT REPEAT 4785 4825 ANK 134.
 FT REPEAT 4821 4861 ANK 135.
 FT REPEAT 4857 4897 ANK 136.
 FT REPEAT 4893 4933 ANK 137.
 FT REPEAT 4929 4969 ANK 138.
 FT REPEAT 4965 5005 ANK 139.
 FT REPEAT 5001 5041 ANK 140.
 FT REPEAT 5037 5077 ANK 141.
 FT REPEAT 5073 5113 ANK 142.
 FT REPEAT 5109 5149 ANK 143.
 FT REPEAT 5145 5185 ANK 144.
 FT REPEAT 5181 5221 ANK 145.
 FT REPEAT 5217 5257 ANK 146.
 FT REPEAT 5253 5293 ANK 147.
 FT REPEAT 5289 5329 ANK 148.
 FT REPEAT 5325 5365 ANK 149.
 FT REPEAT 5361 5401 ANK 150.
 FT REPEAT 5397 5437 ANK 151.
 FT REPEAT 5433 5473 ANK 152.
 FT REPEAT 5469 5509 ANK 153.
 FT REPEAT 5505 5545 ANK 154.
 FT REPEAT 5541 5581 ANK 155.
 FT REPEAT 5577 5617 ANK 156.
 FT REPEAT 5613 5653 ANK 157.
 FT REPEAT 5649 5689 ANK 158.
 FT REPEAT 5685 5725 ANK 159.
 FT REPEAT 5721 5761 ANK 160.
 FT REPEAT 5757 5797 ANK 161.
 FT REPEAT 5793 5833 ANK 162.
 FT REPEAT 5829 5869 ANK 163.
 FT REPEAT 5865 5905 ANK 164.
 FT REPEAT 5901 5941 ANK 165.
 FT REPEAT 5937 5977 ANK 166.
 FT REPEAT 5973 6013 ANK 167.
 FT REPEAT 6009 6049 ANK 168.
 FT REPEAT 6045 6085 ANK 169.
 FT REPEAT 6081 6121 ANK 170.
 FT REPEAT 6117 6157 ANK 171.
 FT REPEAT 6153 6193 ANK 172.
 FT REPEAT 6189 6229 ANK 173.
 FT REPEAT 6225 6265 ANK 174.
 FT REPEAT 6261 6301 ANK 175.
 FT REPEAT 6297 6337 ANK 176.
 FT REPEAT 6333 6373 ANK 177.
 FT REPEAT 6369 6409 ANK 178.
 FT REPEAT 6405 6445 ANK 179.
 FT REPEAT 6441 6481 ANK 180.
 FT REPEAT 6477 6517 ANK 181.
 FT REPEAT 6513 6553 ANK 182.
 FT REPEAT 6549 6589 ANK 183.
 FT REPEAT 6585 6625 ANK 184.
 FT REPEAT 6621 6661 ANK 185.
 FT REPEAT 6657 6697 ANK 186.
 FT REPEAT 6693 6733 ANK 187.
 FT REPEAT 6729 6769 ANK 188.
 FT REPEAT 6765 6805 ANK 189.
 FT REPEAT 6801 6841 ANK 190.
 FT REPEAT 6837 6877 ANK 191.
 FT REPEAT 6873 6913 ANK 192.
 FT REPEAT 6909 6949 ANK 193.
 FT REPEAT 6945 6985 ANK 194.
 FT REPEAT 6981 7021 ANK 195.
 FT REPEAT 7017 7057 ANK 196.
 FT REPEAT 7053 7093 ANK 197.
 FT REPEAT 7089 7129 ANK 198.
 FT REPEAT 7125 7165 ANK 199.
 FT REPEAT 7161 7201 ANK 200.
 FT REPEAT 7197 7237 ANK 201.
 FT REPEAT 7233 7273 ANK 202.
 FT REPEAT 7269 7309 ANK 203.
 FT REPEAT 7305 7345 ANK 204.
 FT REPEAT 7341 7381 ANK 205.
 FT REPEAT 7377 7417 ANK 206.
 FT REPEAT 7413 7453 ANK 207.
 FT REPEAT 7449 7489 ANK 208.
 FT REPEAT 7485 7525 ANK 209.
 FT REPEAT 7521 7561 ANK 210.
 FT REPEAT 7557 7597 ANK 211.
 FT REPEAT 7593 7633 ANK 212.
 FT REPEAT 7629 7669 ANK 213.
 FT REPEAT 7665 7705 ANK 214.
 FT REPEAT 7701 7741 ANK 215.
 FT REPEAT 7737 7777 ANK 216.
 FT REPEAT 7773 7813 ANK 217.
 FT REPEAT 7809 7849 ANK 218.
 FT REPEAT 7845 7885 ANK 219.
 FT REPEAT 7881 7921 ANK 220.
 FT REPEAT 7917 7957 ANK 221.
 FT REPEAT 7953 7993 ANK 222.
 FT REPEAT 7989 8029 ANK 223.
 FT REPEAT 8025 8065 ANK 224.
 FT REPEAT 8061 8101 ANK 225.
 FT REPEAT 8097 8137 ANK 226.
 FT REPEAT 8133 8173 ANK 227.
 FT REPEAT 8169 8209 ANK 228.
 FT REPEAT 8205 8245 ANK 229.
 FT REPEAT 8241 8281 ANK 230.
 FT REPEAT 8277 8317 ANK 231.
 FT REPEAT 8313 8353 ANK 232.
 FT REPEAT 8349 8389 ANK 233.
 FT REPEAT 8385 8425 ANK 234.
 FT REPEAT 8421 8461 ANK 235.
 FT REPEAT 8457 8497 ANK 236.
 FT REPEAT 8493 8533 ANK 237.
 FT REPEAT 8529 8569 ANK 238.
 FT REPEAT 8565 8605 ANK 239.
 FT REPEAT 8601 8641 ANK 240.
 FT REPEAT 8637 8677 ANK 241.
 FT REPEAT 8673 8713 ANK 242.
 FT REPEAT 8709 8749 ANK 243.
 FT REPEAT 8745 8785 ANK 244.
 FT REPEAT 8781 8821 ANK 245.
 FT REPEAT 8817 8857 ANK 246.
 FT REPEAT 8853 8893 ANK 247.
 FT REPEAT 8889 8929 ANK 248.
 FT REPEAT 8925 8965 ANK 249.
 FT REPEAT 8961 9001 ANK 250.
 FT REPEAT 8997 9037 ANK 251.
 FT REPEAT 9033 9073 ANK 252.
 FT REPEAT 9069 9109 ANK 253.
 FT REPEAT 9105 9145 ANK 254.
 FT REPEAT 9141 9181 ANK 255.
 FT REPEAT 9177 9217 ANK 256.
 FT REPEAT 9213 9253 ANK 257.
 FT REPEAT 9249 9289 ANK 258.
 FT REPEAT 9285 9325 ANK 259.
 FT REPEAT 9321 9361 ANK 260.
 FT REPEAT 9357 9397 ANK 261.
 FT REPEAT 9393 9433 ANK 262.
 FT REPEAT 9429 9469 ANK 263.
 FT REPEAT 9465 9505 ANK 264.
 FT REPEAT 9501 9541 ANK 265.
 FT REPEAT 9537 9577 ANK 266.
 FT REPEAT 9573 9613 ANK 267.
 FT REPEAT 9609 9649 ANK 268.
 FT REPEAT 9645 9685 ANK 269.
 FT REPEAT 9681 9721 ANK 270.
 FT REPEAT 9717 9757 ANK 271.
 FT REPEAT 9753 9793 ANK 272.
 FT REPEAT 9789 9829 ANK 273.
 FT REPEAT 9825 9865 ANK 274.
 FT REPEAT 9861 9901 ANK 275.
 FT REPEAT 9897 9937 ANK 276.
 FT REPEAT 9933 9973 ANK 277.
 FT REPEAT 9969 10009 ANK 278.
 FT REPEAT 10005 10045 ANK 279.
 FT REPEAT 10041 10081 ANK 280.
 FT REPEAT 10077 10117 ANK 281.
 FT REPEAT 10113 10153 ANK 282.
 FT REPEAT 10149 10189 ANK 283.
 FT REPEAT 10185 10225 ANK 284.
 FT REPEAT 10221 10261 ANK 285.
 FT REPEAT 10257 10297 ANK 286.
 FT REPEAT 10293 10333 ANK 287.
 FT REPEAT 10329 10369 ANK 288.
 FT REPEAT 10365 10405 ANK 289.
 FT REPEAT 10401 10441 ANK 290.
 FT REPEAT 10437 10477 ANK 291.
 FT REPEAT 10473 10513 ANK 292.
 FT REPEAT 10509 10549 ANK 293.
 FT REPEAT 10545 10585 ANK 294.
 FT REPEAT 10581 10621 ANK 295.
 FT REPEAT 10617 10657 ANK 296.
 FT REPEAT 10653 10693 ANK 297.
 FT REPEAT 10689 10729 ANK 298.
 FT REPEAT 10725 10765 ANK 299.
 FT REPEAT 10761 10801 ANK 300.
 FT REPEAT 10797 10837 ANK 301.
 FT REPEAT 10833 10873 ANK 302.
 FT REPEAT 10869 10909 ANK 303.
 FT REPEAT 10905 10945 ANK 304.
 FT REPEAT 10941 10981 ANK 305.
 FT REPEAT 10977 11017 ANK 306.
 FT REPEAT 11013 11053 ANK 307.
 FT REPEAT 11049 11089 ANK 308.
 FT REPEAT 11085 11125 ANK 309.
 FT REPEAT 11121 11161 ANK 310.
 FT REPEAT 11157 11197 ANK 311.
 FT REPEAT 11193 11233 ANK 312.
 FT REPEAT 11229 11269 ANK 313.
 FT REPEAT 11265 11305 ANK 314.
 FT REPEAT 11301 11341 ANK 315.
 FT REPEAT 11337 11377 ANK 316.
 FT REPEAT 11373 11413 ANK 317.
 FT REPEAT 11409 11449 ANK 318.
 FT REPEAT 11445 11485 ANK 319.
 FT REPEAT 11481 11521 ANK 320.
 FT REPEAT 11517 11557 ANK 321.
 FT REPEAT 11553 11593 ANK 322.
 FT REPEAT 11589 11629 ANK 323.
 FT REPEAT 11625 11665 ANK 324.
 FT REPEAT 11661 11701 ANK 325.
 FT REPEAT 11697 11737 ANK 326.
 FT REPEAT 11733 11773 ANK 327.
 FT REPEAT 11769 11809 ANK 328.
 FT REPEAT 11805 11845 ANK 329.
 FT REPEAT 11841 11881 ANK 330.
 FT REPEAT 11877 11917 ANK 331.
 FT REPEAT 11913 11953 ANK 332.
 FT REPEAT 11949 11989 ANK 333.
 FT REPEAT 11985 12025 ANK 334.
 FT REPEAT 12021 12061 ANK 335.
 FT REPEAT 12057 12097 ANK 336.
 FT REPEAT 12093 12133 ANK 337.
 FT REPEAT 12129 12169 ANK 338.
 FT REPEAT 12165 12205 ANK 339.
 FT REPEAT 12201 12241 ANK 340.
 FT REPEAT 12237 12277 ANK 341.
 FT REPEAT 12273 12313 ANK 342.
 FT REPEAT 12309 12349 ANK 343.
 FT REPEAT 12345 12385 ANK 344.
 FT REPEAT 12381 12421 ANK 345.
 FT REPEAT 12417 12457 ANK 346.
 FT REPEAT 12453 12493 ANK 347.
 FT REPEAT 12489 12529 ANK 348.
 FT REPEAT 12525 12565 ANK 349.
 FT REPEAT 12561 12601 ANK 350.
 FT REPEAT 12597 12637 ANK 351.
 FT REPEAT 12633 12673 ANK 352.
 FT REPEAT 12669 12709 ANK 353.
 FT REPEAT 12705 12745 ANK 354.
 FT REPEAT 12741 12781 ANK 355.
 FT REPEAT 12777 12817 ANK 356.
 FT REPEAT 12813 12853 ANK 357.
 FT REPEAT 12849 12889 ANK 358.
 FT REPEAT 12885 12925 ANK 359.
 FT REPEAT 12921 12961 ANK 360.
 FT REPEAT 12957 12997 ANK 361.
 FT REPEAT 12993 13033 ANK 362.
 FT REPEAT 13029 13069 ANK 363.
 FT REPEAT 13065 13105 ANK 364.
 FT REPEAT 13101 13141 ANK 365.
 FT REPEAT 13137 13177 ANK 366.
 FT REPEAT 13173 13213 ANK 367.
 FT REPEAT 13209 13249 ANK 368.
 FT REPEAT 13245 13285 ANK 369.
 FT REPEAT 13281 13321 ANK 370.
 FT REPEAT 13317 13357 ANK 371.
 FT REPEAT 13353 13393 ANK 372.
 FT REPEAT 13389 13429 ANK 373.
 FT REPEAT 13425 13465 ANK 374.
 FT REPEAT 13461 13501 ANK 375.
 FT REPEAT 13497 13537 ANK 376.
 FT REPEAT 13533 13573 ANK 377.
 FT REPEAT 13569 13609 ANK 378.
 FT REPEAT 13605 13645 ANK 379.
 FT REPEAT 13641 13681 ANK 380.
 FT REPEAT 13677 13717 ANK 381.
 FT REPEAT 13713 13753 ANK 382.
 FT REPEAT 13749 13789 ANK 383.
 FT REPEAT 13785 13825 ANK 384.
 FT REPEAT 13821 13861 ANK 385.
 FT REPEAT 13857 13897 ANK 386.
 FT REPEAT 13893 13933 ANK 387.
 FT REPEAT 13929 13969 ANK 388.
 FT REPEAT 13965 14005 ANK 389.
 FT REPEAT 14001 14041 ANK 390.
 FT REPEAT 14037 14077 ANK 391.
 FT REPEAT 14073 14113 ANK 392.
 FT REPEAT 14109 14149 ANK 393.
 FT REPEAT 14145 14185 ANK 394.
 FT REPEAT 14181 14221 ANK 395.
 FT REPEAT 14217 14257 ANK 396.
 FT REPEAT 14253 14293 ANK 397.
 FT REPEAT 14289 14329 ANK 398.
 FT REPEAT 14325 14365 ANK 399.
 FT REPEAT 14361 14401 ANK 400.
 FT REPEAT 14397 14437 ANK 401.
 FT REPEAT 14433 14473 ANK 402.
 FT REPEAT 14469 14509 ANK 403.
 FT REPEAT 14505 14545 ANK 404.
 FT REPEAT 14541 14581 ANK 405.
 FT REPEAT 14577 14617 ANK 406.
 FT REPEAT 14613 14653 ANK 407.
 FT REPEAT 14649 14689 ANK 408.
 FT REPEAT 14685 14725 ANK 409.
 FT REPEAT 14721 14761 ANK 410.
 FT REPEAT 14757 14797 ANK 411.
 FT REPEAT 14793 14833 ANK 412.
 FT REPEAT 14829 14869 ANK 413.
 FT REPEAT 14865 14905 ANK 414.
 FT REPEAT 14901 14941 ANK 415.
 FT REPEAT 14937 14977 ANK 416.
 FT REPEAT 14973 15013 ANK 417.
 FT REPEAT 15009 15049 ANK 418.
 FT REPEAT 15045 15085 ANK 419.
 FT REPEAT 15081 15121 ANK 420.
 FT REPEAT 15117 15157 ANK 421.
 FT REPEAT 15153 15193 ANK 422.
 FT REPEAT 15189 15229 ANK 423.
 FT REPEAT 15225 15265 ANK 424.
 FT REPEAT 15261 15301 ANK 425.
 FT REPEAT 15297 15337 ANK 426.
 FT REPEAT 15333 15373 ANK 427.
 FT REPEAT 15369 15409 ANK 428.
 FT REPEAT 15405 15445 ANK 429.
 FT REPEAT 15441 15481 ANK 430.
 FT REPEAT 15477 15517 ANK 431.
 FT REPEAT 15513 15553 ANK 432.
 FT REPEAT 15549 15589 ANK 433.
 FT REPEAT 15585 15625 ANK 434.
 FT REPEAT 15621 15661 ANK 435.
 FT REPEAT 15657 15697 ANK 436.
 FT REPEAT 15693 15733 ANK 437.
 FT REPEAT 15729 15769 ANK 438.
 FT REPEAT 15765 15805 ANK 439.
 FT REPEAT 15801 15841 ANK 440.
 FT REPEAT 15837 15877 ANK 441.
 FT REPEAT 15873 15913 ANK 442.
 FT REPEAT 15909 15949 ANK 443.
 FT REPEAT 15945 15985 ANK 444.
 FT REPEAT 15981 16021 ANK 445.
 FT REPEAT 16017 16057 ANK 446.
 FT REPEAT 16053 16093 ANK 447.
 FT REPEAT 16089 16129 ANK 448.
 FT REPEAT 16125 16165 ANK 449.
 FT REPEAT 16161 16201 ANK 450.
 FT REPEAT 16197 16237 ANK 451.
 FT REPEAT 16233 16273 ANK 452.
 FT REPEAT 16269 16309 ANK 453.
 FT REPEAT 16305 16345 ANK 454.
 FT REPEAT 16341 16381 ANK 455.
 FT REPEAT 16377 16417 ANK 456.
 FT REPEAT 16413 16453 ANK 457.
 FT REPEAT 16449 16489 ANK 458.
 FT REPEAT 16485 16525 ANK 459.
 FT REPEAT 16521 16561 ANK 460.
 FT REPEAT 16557 16597 ANK 461.
 FT REPEAT 16593 16633 ANK 462.
 FT REPEAT 16629 16669 ANK 463.
 FT REPEAT 16665 16705 ANK 464.
 FT REPEAT 16701 16741 ANK 465.
 FT REPEAT 16737 16777 ANK 466.
 FT REPEAT 16773 16813 ANK 467.
 FT REPEAT 16809 16849 ANK 468.
 FT REPEAT 16845 16885 ANK 469.
 FT REPEAT 16881 16921 ANK 470.
 FT REPEAT 16917 16957 ANK 471.
 FT REPEAT 16953 16993 ANK 472.
 FT REPEAT 16989 17029 ANK 473.
 FT REPEAT 17025 17065 ANK 474.
 FT REPEAT 17061 17101 ANK 475.
 FT REPEAT 17097 17137 ANK 476.
 FT REPEAT 17133 17173 ANK 477.
 FT REPEAT 17169 17209 ANK 478.
 FT REPEAT 17205 17245 ANK 479.
 FT REPEAT 17241 17281 ANK 480.
 FT REPEAT 17277 17317 ANK 481.
 FT REPEAT 17313 17353 ANK 482.
 FT REPEAT 17349 17389 ANK 483.
 FT REPEAT 17385 17425 ANK 484.
 FT REPEAT 174

DB 1696 GSTAVHPQATSSWQEEVTCGPBHSFORITTT-----OGPPGALQIETGVLYSTREHV 1750

QY 1648 -EGIRETSQVILRPSPNPRTA-----YONENIKSMARKRSORSSYT 1688

DB 1751 QNGPPTGSP--KAGKEPRLMAPESAFSGEVOGDELQNGQVVEEGFT 1798

RESULT 4

ANK3_HUMAN STANDARD; PRT; 4377 AA.

AC 012955:

DT 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DE 15-JUN-2002 (Rel. 41, Last annotation update)

DE Ankyrin 3 (ANK-3) (Ankyrin G).

GN ANK3.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Brain stem;

RX MEDLINE=95138209; PubMed=7836469;

RA Kordeli E., Lambert S., Bennett V.

RT "Ankyrin. A new ankyrin gene with neural-specific isoforms localized at the axonal initial segment and node of Ranvier."

RL J. Biol. Chem. 270:2352-2359(1995).

CC -1- FUNCTION: Membrane-cytoskeleton linker.

CC -1- ALTERNATIVE PRODUCTS: A number of isoforms are produced by alternative splicing.

CC -1- TISSUE SPECIFICITY: Expressed in brain and other tissues.

CC -1- SIMILARITY: CONTAINS 23 ANK REPEATS.

CC -1- SIMILARITY: CONTAINS 1 DEATH DOMAIN.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: U13616; AAA64834.1; -

DR HSP: P53273; 1B18.

DR Genew: HGNC:494; ANK3.

DR MIM: 600465; -

DR InterPro: IPR002110; ANK.

DR InterPro: IPR000488; Death.

DR InterPro: IPR000906; ZUS.

DR Pfam: PF00023; ank; 24.

DR Pfam: PF00531; death; 1.

DR Pfam: PF00791; ZUS; 1.

DR PRINTS: PRO1415; ANKYRIN.

DR SMART: SM00248; ANK; 21.

DR SMART: SM00005; DEATH; 1.

DR SMART: SM00218; ZUS; 1.

DR PROSITE: PS50088; ANK_REPEAT; 21.

DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.

DR PROSITE: PS50017; DEATH_DOMAIN; 1.

KW Cytoskeleton; Alternative splicing; Repeat; ANK repeat.

FT REPEAT 73 102 ANK 1.

FT REPEAT 106 135 ANK 2.

FT REPEAT 139 168 ANK 3.

FT REPEAT 172 201 ANK 4.

FT REPEAT 203 230 ANK 5.

FT REPEAT 234 263 ANK 6.

FT REPEAT 267 296 ANK 7.

FT REPEAT 300 329 ANK 8.

FT REPEAT 333 362 ANK 9.

FT REPEAT 366 395 ANK 10.

FT REPEAT 399 428 ANK 11.

FT REPEAT 432 461 ANK 12.

FT REPEAT 465 494 ANK 13.

FT REPEAT 498 527 ANK 14.

FT REPEAT 531 560 ANK 15.

FT REPEAT 564 593 ANK 16.

FT REPEAT 597 626 ANK 17.

FT REPEAT 630 659 ANK 18.

FT REPEAT 663 692 ANK 19.

FT REPEAT 696 725 ANK 20.

FT REPEAT 729 758 ANK 21.

FT REPEAT 762 791 ANK 22.

FT REPEAT 795 825 ANK 23.

FT DOMAIN 1519 1898 SER-RICH.

FT DOMAIN 4090 4174 DEATH.

FT SEQUENCE 4377 AA; 480399 MW; F42379E55768B684 CRC64;

Query Match 5.9%; Score 526; DB 1; Length 4377;

Best Local Similarity 20.0%; Pred. No. 2,1e-20;

Matches 416; Conservative 266; Mismatches 664; Indels 734; Gaps 80;

QY 14 EEENIPALKALLEKCKVDNEREGQPTPLMLAEGNVEIYKELIKNGANCNLEDLQNT 73

DB 83 KEGHEVVESELLQREANVDAATKGNLTALHTASLAGAEVVKVLTNGANVNAOSQNGFT 142

QY 74 ALISAKEGHIIHVELEKSGASLEHRDGMGTALMAVCYGRDVELL----- 124

DB 143 PLYMAAENHLEVFLLDNGASQSLATEDGFTPLVALQGHQGVSLLENDTKQVR 202

QY 125 -----SHGANPSYTG- 135

DB 203 LPALHTARKDDTKAALLLQDNNAVDESKGFTPLHIAHYGINVATLLNRAAVD 262

QY 136 ---QSYVPIIWAARGHADIYHLLQNGAVKCSDKXTGTPYWAARKG----- 182

DB 263 FTARDITDPLHAVASRGNAMVKLLDNGAKIDATRGTLPLHGASGHEQVEMLLD 322

QY 183 -----HLECYHLLAMGADVDOEGANSMTALVAVKGTYQS 219

DB 323 RAAPLSTKKNKLSPLHMAQGDHLNVOYLLQHNVPYDDVTDYTLALHVAAGHYKV 382

QY 220 VEEILKRNPVNLTKDGTALMISKE----- 247

DB 383 AVLLDKRNPNAKALNGFTPLHACKNRIRKVMELLKHGASIQAVTESGLTPHVAF 442

QY 248 -GHIEIYODLLDAGTYVNPDRSGPTVLIGAVRGHVEIVRALLQKYADIDIRGDNRTA 306

DB 443 MGHVNIQSOLHHRGASPTNTTVRGFTALHMAARSGAEVVRVLYDGDAGVLEKAKDDPT 502

QY 307 LYWAVERGNATVVRDILQ--CNPD- 333

DB 503 LHTSARLGRADIVQOLLQOGASPNNAATSGYTPPLHSAREGHEDVAEFLDHGASLSITT 562

QY 334 KQGETPLIKATKMRNIEVEVELLDKGAKYSAVDKKGDPPLHVAIGSRRLAELLRMPK 393

DB 563 KKGFTPLHVAARYKLEVANLLLOKSASPDAAKSGLPLHVAHYDNQKVALLL--D 619

QY 394 DGRILYRPKAGETPNIDC-SHOKSILTOI--FGAR-----HSPETDGD 438

DB 620 QGASPHAAAKNGTTPPLHIAKKNOADIATTLLEYGDANAAVYTRQGIASVHLAAGSHVD 679

QY 439 LGYDISSALADILSEPTMQ-----PVCGLY 466

DB 680 VSLILGRNANVNLSKSGLPLHLAAGEDRVNVAEVLVNOGAHVDAQTKMGYTPLVGCH 739

QY 467 AOWGSK--SFLIK-KLEDEMKTFGQGTPEPLFQSWL--YVFLTLLOGGLGVFAFP- 520

DB 740 --YGNIKIVPFLQSHAKVNAKTKNGY--TPLHQAOQCHTIIIVLQNN-----ASPN 790

QY 521 ---VDTNLAIAISLSFLALIIYIFVIYVGRREGESNMAMALSTRALHIGYLELLEK 577

DB 791 ELTVNNTALGI-----ARRIGIYSVD 814

QY 578 LMFV-----NPPE-----LPEQTTKALPVRFLEFTYNNRSSV--GGFT 613

```

Db 815 LKIVTEETTTTTEKHKMNPETMNEVLMDSDDEVKANAPEMLSGEYISDVEGED 874
Qy 614 SIA-----EMIALSDACEREFGLATRL-----FRYFRRESOGKKKKKCTCL 658
Db 875 ANTGDTDKLGPQDLKELCDDSLPAEGTNGFSLGKRSLSFSDSRYTTLN--RSSAR 932
Qy 659 PSFVIFLIVGCIITGILLALFRVP- KHLT-----VNALISIASVGLAFVINC 709
Db 933 DSMMEILLVPSKEOHLTFTRREFDSDSLRHYSMADTLDNVNLVSSPIHSGFLVSFWDA 992
Qy 710 RTWQVVLDSILNSOKRRLHSAASKLHKLSECFMVL-----KCEVELMARAKITDSTQ 765
Db 993 RG-----GSMGSRHHC-----MRITPPKCTAP-----TR 1019
Qy 766 NOTRLVITIDGADCEODKVLQMLDTVRLFSKGFIAFASDPHIIKAINONLSYL 825
Db 1020 ITCRLV-----KRHKLAN-----PPH-----GERRGISRLV 1047
Qy 826 DSNINGHYMRN-IYHLFVFLNSGLSNARKFLVTSATNGDITGSDTTGTQEDTDRVSQ 884
Db 1048 EMGPAGAOFLGPVIVEIPHFGSMRG--KERELIVLRSENGE-----TWREHOFD--SK 1096
Qy 885 NSLGEMTKL--GSKTALNRDQYRRRQQRITQMSFDLKLVLTEDEWESDISQIMRR 942
Db 1097 NE--DITELNGMDELDSPELGKKRICKRITITDF-----POYFAVS--RIKQ 1142
Qy 943 LLNIVSVTGRLLRAN-----QITFMWRSLASMINLTEQWPRYSMLIYLETGLPDPM 997
Db 1143 ESNQIGPBGILSTFVPLVQASPEGALTRKINVLGQ-----APVDEI 1188
Qy 998 TLKTN-----YERISKNP-----TTKDEPLEI 1022
Db 1189 VKKILGNKATSPVIVYVPRRRKFKPTMTIPIVPPSGEGSVNGYKDTPNRLCSII 1248
Qy 1023 DG-----DIRNEFVLSRTPVLVARDVKTFPCVNDPLKRLIADVRAARQIN 1074
Db 1249 TGGSPAQMEDT-----TGTTLTFLTKDCVSF--TTNV--SARWMLADCHQVLETVG 1296
Qy 1075 IGGIAYPPLPLHEGPPPPPGSYSOPASVCSASFNGPPGGVSPDHSSYYSGLSGPOH 1134
Db 1297 LATQLYREL-----ICV-----PYMAKFV-----VEAKMNP-- 1323
Qy 1135 PFYRAAVPATGSSILLSSMTVDVVCCKRIQIEGLDNNMPOYCTTIKKANINRVLSQC 1194
Db 1324 -----VSSLRCPMTDDVDVTLDEQE----- 1346
Qy 1195 NIDELKKEMANFGDMLFRSMVLEKRSVSOVVPEDPRFLN-ENSSAPVPHGE-----S 1248
Db 1347 NFEV-----ARSKOIEVLEGRPIYVDCGNLAPLTKGQQLVEN 1386
Qy 1249 ARBSHTLP-----LTELSSQTPYTLNFSFEELNTLGLDCAPRHSLMSQOTRTPSL 1304
Db 1387 FYSFKENRLPFSIKIRTSQPCGRISFLKERKTKTGLPOTAVCNLITLPAHKKETES- 1445
Qy 1305 SLSNSQDSSTFISKLTQKVOAEYDAYRE--YIAQMSQLEGGGSSSTIGRSSPHSTY 1361
Db 1446 -----DODDEIF--KTRRQSPASLALRKRYSTLTCMTERTSGA--RSLP--TTY- 1492
Qy 1362 IGOSSSGGSIHSTLEOGRGEGELKODGKRSFLMKGDVIDYSSSGSVSTNEASPLDPT 1421
Db 1493 -----SYKPFSTPYQSWTAPITVGPAS-----GPTSLSSSSNTPSPASPLKSTW 1541
Qy 1422 EDEKSNQSGSKLLPGKSSERPSLPOTDLKLGGLRQKLPDSDEDESGTGVQITPHC 1481
Db 1542 -----SVSTSPSPKSTL-----GASTTSVKSISDV 1567
Qy 1482 SKMIRTKRLAKORECASPEHSAEPIRTFIRAKY-----LSDALLDK-----KXS 1528
Db 1568 ASPIRSLRT-----MSPKIYVSSPINIOVSSGTLAPAPVATPLKGI 1614
Qy 1529 SDGCVRSNESSPMHSLHNEAADSQLEKANLIELEDEGHGSKRGMPHS----- 1576

```

```

Db 1615 ASNTESSRTSP-----VTTAGSLERSSTITMTPPASPKSNINMTSSSLPFKSIITSA 1668
Qy 1577 -----LSGIQDPIAMISICSEDKSPSCSLIASSPERSWA----- 1614
Db 1669 PLISSPLKSVSPKSVSHVDIVISSAKITMASS-----LSSPVQMGHAEVALVNGISPLKY 1725
Qy 1615 COKAYNNRTPSTVTLNKN--TAPUNRANQ-----NFEIEIGIRETSQVI-----LRPSPN 1665
Db 1726 ASSTLNGKATATLQEKATISSATNSVSVSAATQVEVSTTAMPSPPLKSTVSA 1785
Qy 1666 PTAVQENLKSMAHKSQSSYTRLSKDSSELHAASSEST 1705
Db 1786 PSAFOSLRTPSPAS-----ALYTSL--GSSISATTSVYT 1816

RESULT 5
YB23_HUMAN
ID YB23_HUMAN STANDARD; PRT; 768 AA.
AC Q9ULJ7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein KIAA1223 (fragment).
GN KIAA1223.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MDLINE=20039619; PubMed=10574462;
RA Nagase T., Ishikawa K.-I., Kikuno R., Hirose M., Nomura N.,
RA Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XV.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 6:337-345(1999).
CC - SIMILARITY: CONTAINS AT LEAST 14 ANK REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
DR EMBL: AB033049; BAA86537.1; -.
DR HSSP: P42771; IDC2.
DR InterPro: IPR002110; ANK.
DR Pfam: PF00023; ank. 14.
DR SMART: SM00248; ANK. 13.
DR PROSITE: PS50088; ANK_REPEAT. 13.
DR PROSITE: PS50297; ANK_REPEAT_REGION. 1.
KM Hypothetical Protein; Repeat; ANK repeat.
FT NON_TER 1 1
FT REPEAT 1 11 ANK 1.
FT REPEAT 15 44 ANK 2.
FT REPEAT 48 82 ANK 3.
FT REPEAT 86 115 ANK 4.
FT REPEAT 119 148 ANK 5.
FT REPEAT 152 181 ANK 6.
FT REPEAT 185 214 ANK 7.
FT REPEAT 218 247 ANK 8.
FT REPEAT 251 280 ANK 9.
FT REPEAT 284 313 ANK 10.
FT REPEAT 317 346 ANK 11.
FT REPEAT 350 379 ANK 12.
FT REPEAT 383 412 ANK 13.
FT REPEAT 416 446 ANK 14.
FT SEQUENCE 768 AA; 82819 MW; 2913B69BE2DFE06D CMC64;

```


Db 490 DSGKCLEYLLRNDANPGI---RDKQGNVAVSAAYGHRLQLDLA-----SEPLDYLM 540

OY 441 YDLSSALADILSEPMOPPIGVGLAOWSGSKFLKLED-EKKTAGQOTEPLOFS 499

Db 541 ETSGIDMSDSRRATIS-PLHLAAVHGHQALEVLYVSLDLVNRSSGRTPLDLAER 599

OY 500 WLIVPLTLLICGSLAV 516

Db 600 GHVECVDLINOGASIL 616

RESULT 7

DAER_HUMAN

ID DAPK_HUMAN STANDARD; PRT; 1431 AA.

AC P53355;

DT 01-OCT-1996 (rel. 34, Created)

DT 01-NOV-1997 (rel. 35, Last sequence update)

DT 16-OCT-2001 (rel. 40, Last annotation update)

DE Death-associated protein kinase 1 (EC 2.7.1.-) (DAP kinase 1).

GN DAPK1 OR DAPK.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=95129831; PubMed=7828849;

RA Deiss L.P., Feinstein E., Berissi H., Cohen O., Kimchi A.

RT Identification of a novel serine/threonine kinase and a novel 15-kD protein as potential mediators of the gamma interferon-induced cell death.

RT genes Dev. 9:15-30(1995).

RL [2]

RN REVISIONS TO 164-171.

RA Feinstein E.

RT Submitted (Apr-1997) to the EMBL/Genbank/DBJ databases.

CC -1- FUNCTION: INVOLVED IN MEDIATING INTERFERON-GAMMA-INDUCED CELL DEATH.

CC -1- PFM: AUTOPHOSPHORYLATED.

CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

CC -1- SIMILARITY: CONTAINS 10 ANK REPEATS.

CC -1- SIMILARITY: CONTAINS 1 DEATH DOMAIN.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

CC -----

EMBL: X76104; CA53712.1; -

HSSP: Q63450; IAO6.

DR Genew: HGNC:2674; DAPK1.

DR MIM: 600831; -

DR InterPro: IPR002110; ANK.

DR InterPro: IPR000488; Death.

DR InterPro: IPR000719; Euk_pkinase.

DR InterPro: IPR002290; Ser_thr_pkinase.

DR Pfam: PF00023; ank; 8.

DR Pfam: PF00069; pkinase; 1.

DR Pfam: PF00531; death; 1.

DR ProDom: PRO00001; Euk_pkinase; 1.

DR SMART: SM00248; ANK; 7.

DR SMART: SM00005; DEATH; 1.

DR SMART: SM00220; S_TKC; 1.

DR PROSITE: PS50088; ANK_REPEAT; 6.

DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.

DR PROSITE: PS50017; DEATH_DOMAIN; 1.

DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.

DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.

DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.

KW Transferase; Serine/threonine-protein kinase; Calmodulin-binding;

KW Phosphorylation; ATP-binding; Repeat; ANK repeat; Apoptosis.

FT DOMAIN 13 266 PROTEIN KINASE.

FT DOMAIN 267 334 CALMODULIN-BINDING.

FT REPEAT 378 407 ANK 1.

FT REPEAT 411 440 ANK 2.

FT REPEAT 444 473 ANK 3.

FT REPEAT 478 507 ANK 4.

FT REPEAT 511 540 ANK 5.

FT REPEAT 544 573 ANK 6.

FT REPEAT 577 606 ANK 7.

FT REPEAT 610 639 ANK 8.

FT REPEAT 876 905 ANK 9.

FT REPEAT 1163 1197 ANK 10.

FT NP_BIND 1313 1397 DEATH.

FT BINDING 19 27 ATP (BY SIMILARITY).

FT ACT_SITE 42 42 ATP (BY SIMILARITY).

FT MUTAGEN 139 139 BY SIMILARITY.

FT SEQUENCE 1431 AA; 160017 MW; 9EB84811004A1558 CRC64;

Query Match 4.78; Score 414.5; DB 1; Length 1431;

Best Local Similarity 21.68; Pred. No. 5.5e-15;

Matches 221; Conservative 156; Mismatches 335; Indels 309; Gaps 36;

OY 5 ISOSVINYVEEENIPALKALECK--DVERNECGOTPLMLAEQGNVEIYELKNGA 62

Db 344 YMKRITHAINDVNPGLIHLGLSLSDVNPQPKHGTPLPLIAGCGNLIQLIKRGS 403

OY 63 NCNLEDINTALISAKESGHIHVEELKSGASLEHRDGMGTALMAMCYGRDYLVL 122

Db 404 RIDVDKGSNNAYMAARHGHVDTLKFLSENKCPDLVDKSGEMALHVARGHADVAV 463

OY 123 LLSHGAMPSTGYQSYVPIIIMAGRGHADIVILLQNGAKVNSDKYGTPLVMAARG 182

Db 464 TCASAPIPISRKKEETPLHCAMHGYSVAKALCEACCNVNIKRKEETPLTASAG 523

OY 183 HLBCVKKLMDGADVQEGANSMTALIVAKGGYTSYKEILKRNPNVMTDKDNTALM 242

Db 524 YHDIVECLAHGAD-----LNACCKDGHIALH 550

OY 243 IASKEGHEIYVOLLADAGTYVNIIPDSGDTLVIGARGHVEIVRALLDKYADIDRQD 302

Db 551 LAVRRCOMEYIKTLISQGCYVDYDRHGNTPHVAACKDNMPYVALCE----- 599

OY 303 NKTALYAAVEKGNATWRILOCPNTEICTKDGEPILKATKRMRIEIVELLDKAVY 362

Db 600 -----ANCNLD--ISKNGRPLHLAANGILDVYRLCLMGASY 637

OY 363 SAVDKKGDPLHVAIRGSRRLAELLRNPKDGRLLYRPNKAGETPYNIDCSHOKSTLQ 422

Db 638 EALTGDKTAEDLARSEQHENVAGLLARLKD-----THRGLEIQQ 678

OY 423 IFGARHLSPEITDGMGLGVDIVSSALADILSEPTMQPPICVGLYAKOWSGSKFLKLED 482

Db 679 -----LRPQQ-----NLOPRIKILFHHSGSGKTTLVESLK- 709

OY 483 EMKTFAGQOTEPLOFSWLIVETLLICGGLV-----AFVDYTLA 526

Db 710 -----CGLLASFFRRRRRLSTSNSSFFPPSLASKPT 742

OY 527 IATLSFLALIYFFIVITVYGGRRGEGSNMMAMALSTRLARHIGVLELLFKIMFVNPPEL 586

Db 743 VSVSINNLT-----YPCGENSVYRSRSMPEPGLTK--GMLE-----VFVAPTHH 784

OY 587 P-----EOTTKAL-----PYRPLFTDYNKLSSVSGGTSLAEMLA 620

Db 785 PHCSADQSTKAIDIONAVLNVGDFSWERSGNPNVFCYDY-----FANDFTSHVAVF 841

OY 621 TISDACERFEG---FLATRLFRVFRTES---QGRKKMKTKCLPSYIFLFIYGCILAG 674

Db 842 SLEEPYEQLNVPVFWLFLSKLSLVVEEPIAGGRLK-----NPLQYVL----- 885

OY 675 ITLLAIFRVDPKHLVYNAIILISIASVGLAFVLANCRITWQVYDLSLNSQRRK---LHSA 730

Db 886 -----VATHADIMNPRAGEFGDKDT-----SLIKETRNFGNDLH-I 925
 QY 731 ASKRLHKLSEGF---MKYLKCEVELMARMAKTIDSPONQRLVYIIDGLACEODKVL 786
 Db 926 SNKLFVLIDGASGSKDMKVLRLNHLQ---EIRSOIVSVCPRMHL-----CE--KII 971
 QY 787 QMLDIYRVLFKSPFPFALF--FASDPHIIIIKATMONNSVLRSNGHDMRYIVLPLVF 844
 Db 972 STLPSPWRKLNQPNQMLSLQOFYVD-----VODQINPLASEED-----LRRIAQO--- 1015
 QY 845 LNSRGSLNARKFLVTSATNDICSDPTGTHQEDTDNRVSONSLGEMTKLSKLTALNR-RD 903
 Db 1016 LHETGEIN-----IMQSEIVQDVLLD-----PMLCTNVLGKLSLSETPRALHMYRG 1063
 QY 904 TYRRQWORTITRQMSFDTKLKLVTEWMS-DISPOTRMLNIVSVTGLRLANDITFN 962
 Db 1064 RYVEDIQRLVPSDVEELLQIILDAMDICARDLSGTM-----VDVPA-LIKTDMLHRS 1116
 QY 963 W 963
 Db 1117 W 1117

RESULT 8

TNKL_HUMAN STANDARD; PRT; 1327 AA.
 ID TNKL_HUMAN
 AC 095271; 095272;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE Tankyrase 1 (EC 2.4.2.30) (TANK1) (Tankyrase I) (TNS-1) (TRF1-
 GN interacting ankyrin-related ADP-ribose polymerase).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NX NCBI_Taxid=9606;
 RX [1]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RC TISSUE-Testis;
 RX MEDLINE=99040105; PubMed=9822378;
 RA Smith S., Giliat I., Schmitt A., de Lange T.;
 RT "Tankyrase, a poly(ADP-ribose) polymerase at human telomeres.";
 RL Science 282:1484-1487(1998).
 RN [2]
 RP SUBCELLULAR LOCALIZATION.
 RX MEDLINE=99454782; PubMed=10523501;
 RA Smith S., de Lange T.;
 RT "Cell cycle dependent localization of the telomeric PARP, tankyrase,
 RL to nuclear pore complexes and centrosomes.";
 RN J. Cell Sci. 112:3649-3656(1999).
 RN [3]
 RP FUNCTION, AND PHOSPHORYLATION.
 RX MEDLINE=20556282; PubMed=10988299;
 RA Chi N.-W., Lodish H.F.;
 RT "Tankyrase is a golgi-associated mitogen-activated protein kinase
 RL substrate that interacts with IRAP in GLUT4 vesicles.";
 RN J. Biol. Chem. 275:38457-38444(2000).
 RN [4]
 RP FUNCTION, AND MUTAGENESIS OF HIS-1184 AND GLU-1291.
 RX MEDLINE=21602874; PubMed=11739745;
 RA Cook B.D., Dynek J.N., Chang W., Shostak G., Smith S.;
 RT "Role for the related poly(ADP-ribose) polymerases tankyrase 1 and 2
 RL at human telomeres.";
 RN Mol. Cell. Biol. 22:332-342(2002).
 CC -1- FUNCTION: May regulate vesicle trafficking and modulate the
 CC subcellular distribution of SLC24A4/GLUT4-vesicles. Has PARP
 CC activity and can modify TRF1, and thereby contribute to the
 CC regulation of telomere length.
 CC -1- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribose]1(N)-acceptor =
 CC nicotinamide + [ADP-D-ribose]1(N+1)-acceptor.
 CC -1- SUBUNIT: Oligomerizes and associates with TNS2. Interacts with

CC the cytoplasmic domain of LNP/OTase in SLC24A4/GLUT4-vesicles.
 CC Binds to the N-terminus of telomeric TRF1 via the ANK repeats.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic, associated with the Golgi and
 CC with juxtanuclear SLC24A4/GLUT4-vesicles. A minor proportion is
 CC also found at nuclear pore complexes and around the pericentriolar
 CC matrix of mitotic centrosomes. During interphase, a small fraction
 CC of TNS is found in the nucleus, associated with TRF1.
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; may be
 CC produced by alternative splicing.
 CC -1- TISSUE SPECIFICITY: Ubiquitous; highest levels in testis.
 CC -1- PTM: Upon insulin-stimulation, phosphorylated on serine residues
 CC by MAPK kinases.
 CC -1- SIMILARITY: BELONGS TO THE PARP FAMILY.
 CC -1- SIMILARITY: CONTRAINS 15 ANK REPEATS.
 CC -1- SIMILARITY: CONTRAINS 1 SAM DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

CC EMBL: AF082556; AAC79841.1; -;
 DR EMBL: AF082557; AAC79842.1; -;
 DR EMBL: AF082558; AAC79843.1; -;
 DR EMBL: AF082559; AAC79844.1; -;
 DR HSSP: Q00420; IAWC.
 DR GeneW: HGNC:11941; TNS.
 DR MIM: 603303; -;
 DR InterPro: IPR002110; ANK.
 DR InterPro: IPR001660; SAM.
 DR Pfam: PF00023; ank; 21.
 DR PRINTS: PRO1415; ANKYRIN.
 DR SMART: SM00248; ANK; 15.
 DR SMART: SM00454; SAM; 1.
 DR PROSITE: PS50088; ANK_REPEAT; 15.
 DR PROSITE: PS02927; ANK_REPEAT_REGION; 1.
 DR PROSITE: PS0105; SAM_DOMAIN; 1.
 KW Nucleotidyltransferase; Glycosyltransferase; NAD; Golgi stack; Telomere;
 KW Nuclear protein; Repeat; ANK repeat; ADP-ribosylation;
 KW Phosphorylation; Alternative splicing.
 FT REPEAT 215 247 ANK 1.
 FT REPEAT 248 280 ANK 2.
 FT REPEAT 281 313 ANK 3.
 FT REPEAT 368 400 ANK 4.
 FT REPEAT 401 433 ANK 5.
 FT REPEAT 434 466 ANK 6.
 FT REPEAT 521 556 ANK 7.
 FT REPEAT 557 589 ANK 8.
 FT REPEAT 590 622 ANK 9.
 FT REPEAT 683 715 ANK 10.
 FT REPEAT 716 748 ANK 11.
 FT REPEAT 749 781 ANK 12.
 FT REPEAT 836 868 ANK 13.
 FT REPEAT 869 901 ANK 14.
 FT REPEAT 902 934 ANK 15.
 FT REPEAT 1030 1089 SAM.
 FT DOMAIN 9 14 PARP.
 FT DOMAIN 137 145 POLY-HIS.
 FT DOMAIN 128 134 POLY-PRO.
 FT DOMAIN 137 145 POLY-SER.
 FT DOMAIN 137 145 POLY-SER.
 FT VARSPIC 644 643 EST -> GHS (IN ISOFORM 2).
 FT VARSPLIC 644 1327 MISSING (IN ISOFORM 2).
 FT MTAGEN 1184 1184 H->A: LOSS OF ACTIVITY; WHEN ASSOCIATED
 FT MTAGEN 1291 1291 E->A: LOSS OF ACTIVITY; WHEN ASSOCIATED
 FT MTAGEN 1291 1291 WITH A-1184.
 SQ SEQUENCE 1327 AA; 142010 MW; E14DE985C710B957 CRC64;

Query Match 4.5%; Score 395.5; DB 1; Length 1327;
 Best Local Similarity 23.3%; Pred. No. 5.5e-14;
 Matches 167; Conservative 80; Mismatches 211; Indels 259; Gaps 17;

14 EENIPALKALEKCDVERNEGOTPLMLAEGONGEIVKELKNGANCLLEDLDMWT 73
 348 EKKIMALLPLVNVCHASDGRK---STPLHLAAGNRRKRIYVOLLQHGADVAAKDKGLV 404
 74 ALISAKESGHHIIVEELKSGASLEHRDNGMTALMAMCYKRTDVELLSHGANSPV- 132
 405 PLHNCSTGYHEVTELLKHGACVAMDMQFTPLHEAASKRVEVCSILLSHGADPTLV 464
 133 -----T 133
 465 NCHGSAVDMAPPELREKRLTEYEFKSHLLQARADLAKYKKTALLETINFKOPQSHET 524
 134 GLQYSV-----YPLIMAGRGHADIHLLONGAKV 164
 525 ALHCANVASLHPKRVKQVTELLRKGANVNEKNKDFPLHVAERAHNDVMEVLHKGAKM 584
 165 NCSDKYCTPLVMAARKHLECVKLLMAGADVDOEGANSMTA----- 207
 585 NALDTLGGTALRALAHLQTCRLLSYSDPSIISLQGTAAQMGNEAVOOLISESTP 644
 208 -----LIVAVAGYQSYVEIKLRNPVNLTKDG--NTALMISKEGHIEYVDL 256
 645 IRTSDVDRLEBASAGDELEYKOLCS--SQNVNCHDLGRSHSTPLHFAAGYNRSVVEYL 703
 257 LDAGTYVNPDRSGTIVLIGAVRGHVEIVRALLOKVIDIDIRGODNKLTALYMAVEKG- 314
 704 LHHGADVAAKDKGLVPLHNCSTGYHEVTELLKHGACVAMDMQFTPLHEAASKRKY 763
 315 -----NATVNDILO-----C 325
 764 EICKLLKHGADPTKKNNDGNTPLDYKEGDTIDDLKGAALLDAKKCCLARVOKLC 823
 326 NEDTICTKD---GETPLIKATKRNIEVELLDKGAVSADVKKGDTPLH----- 374
 824 TPENINCRDQGRNSTPLHLAAGYNNLEVAEYLLEHGADVANAOKGGLIPLHNAASYGHV 883
 375 -----VAIRGRSRRLALLLRNPDKGLLYRPKAGETP 408
 884 DIALLLKYNTCVNATDWMATPLHEAOKGRT-QLCLALLHAGADPM--KKQEGCTP 939
 409 YN-----IDCSHOKSILT-----OIFGARHLSPTETDGMGLYDYSALADI 451
 940 IDLATADDIRALLIDAMPPEALPTCFKQATVVASLISPASTPS-----CLSAASSTDN 994
 452 LSEPTMOPPIGVGLYAOVSG-----KSFLLKLEDEMKT 486
 995 LTGPLAE--LAVGGSANAGDAAGTERKEGEGVAGLDMNISQFLSLGLEHLDIFET 1049

RESULT 9
 ASB2_HUMAN STANDARD; PRT; 587 AA.
 ID ASB2_HUMAN STANDARD; PRT; 587 AA.
 AC Q96Q27; Q9NSU5; Q9Y567;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Ankyrin repeat and SOCS box containing protein 2 (ASB-2).
 GN ASB2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND INDUCTION.
 RX MEDLINE=21450655; PubMed=1156180;
 RA Kohroki J., Fujita S., Itoh N., Yamada Y., Imai H., Yumoto N.,
 RA Nakanishi T., Tanaka K.;
 RA "ATRA-regulated Asb-2 gene induced in differentiation of HL-60

RT leukemia cells.";
 RL FEBS Lett. 505:223-228(2001).
 RN [12]
 RP SEQUENCE FROM N.A., AND INDUCTION.
 RX MEDLINE=21634875; PubMed=11682484;
 RA Gubal F.C., Moog-Lutz C., Smolewski P., Di Gioia Y.,
 RA Dazyniewicz Z., Lutz P.G., Cayre Y.E.;
 RT "ASB-2 inhibits growth and promotes commitment in myeloid leukemia
 cells.";
 RL J. Biol. Chem. 277:218-224(2002).
 RN [13]
 RP SEQUENCE OF 31-587 FROM N.A.
 RC Tissue-uterus;
 RA Oltensaeider B., Obermaier B., Mewes H.-W., Gassenhuber J.,
 RA Wiemann S.;
 RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.
 RN [14]
 RP SEQUENCE OF 42-587 FROM N.A.
 RX MEDLINE=20564172; PubMed=11111040;
 RA Kile B.T., Vanev E.M., Willison T.A., Brodnicki T.C., Cancilla M.R.,
 RA Herlihy A.S., Croker B.A., Baca M., Nicola N.A., Hilton D.J.,
 RA Alexander W.S.;
 RT "Cloning and characterization of the genes encoding the ankyrin repeat
 and SOCS box-containing proteins Asb-1, Asb-2, Asb-3 and Asb-4.";
 RL Gene 258:31-41(2000).
 CC -1- FUNCTION: Not known. Its expression induces growth inhibition and
 chromatin condensation.
 CC -1- INDUCTION: By all-trans retinoic acid (ATRA).
 CC -1- SIMILARITY: CONTAINS 12 ANK REPEATS.
 CC -1- SIMILARITY: CONTAINS 1 SOCS BOX DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation-
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

EMBL; AB056723; BAB64532.1;
 EMBL; AJ251338; CAC17765.1;
 EMBL; AL137735; CAB70899.1;
 EMBL; AF159164; AAD45345.1;
 GenBank; HGNC:16012; ASB2.
 MIM; 605759;
 InterPro; IPR002110; ANK.
 InterPro; IPR001496; SOCS.
 Pfam; PF00023; ank.11.
 PROSITE; PS50088; ANK_REPEAT. 9.
 PROSITE; PS50297; ANK_REPEAT_REGION. 1.
 PROSITE; PS50225; SOCS. 1.
 ANK repeat; Repeat. 1.
 REPEAT 56 85
 REPEAT 89 119 ANK 1.
 REPEAT 123 152 ANK 2.
 REPEAT 156 185 ANK 3.
 REPEAT 189 218 ANK 4.
 REPEAT 222 251 ANK 5.
 REPEAT 255 284 ANK 6.
 REPEAT 288 317 ANK 7.
 REPEAT 320 349 ANK 8.
 REPEAT 362 391 ANK 9.
 REPEAT 392 421 ANK 10.
 REPEAT 428 456 ANK 11.
 REPEAT 453 587 ANK 12.
 DOMAIN 533 587 SOCS BOX.
 FT 224 224 MISSING (IN REF. 4).
 FT 460 460 MISSING (IN REF. 4).
 FT 460 460 MISSING (IN REF. 4).
 SQ SEQUENCE 587 AA; 65083 MW; 866A2B6010C4C1F9 CRC64;

Query Match 4.3%; Score 384.5; DB 1; Length 587;
 Best Local Similarity 30.4%; Pred. No. 6.2e-14;
 Matches 113; Conservative 71; Mismatches 161; Indels 27; Gaps 6;

QY VINVEBENIPALKALECKCDYDERNEGOTPLMLAEOGNEYIVEELK-NGANCNLE 67
 Db 61 LIRAIXKGDDEALKTKMKRKNLAEPKKEMLPHEAAYGQYGLKLVORALPGITDOR 120
 QY 68 DLDNMLTASAKRGEHIIHEVIELLSGASLEHRDMGWTALMMACYKGRDVELLSHG 127
 Db 121 TLGEFAVYLVATCGHIDCLLSLQAGAEBDISNKRSHETPLKACERKNAEAVKILVOHN 180
 QY 128 A-----NPSVTGQGYVYPIYMAAGRGHADIYHLLQNAKNCSTKCTTPLYMAARK 181
 Db 181 ADTNHRCNRGWTALHESV-----SRNDLEVMOILVSGAKESKNAKYGITPLPEYAQS 233
 QY 182 GHLECVHLLAMGADVQEGANSMTALIVAKGYTOSVEILKRNPNVMLTRKDGNTAL 241
 Db 234 GQLEALFLAKYGDINTFOASDNASALYEACKNHEVEFLLSQGADANKTKNDGLPL 293
 QY 242 MIASKEGHEIYDLDLAGYVNIIPDRSDTVLIGAVRGHVEIVRALLQKYADID- 297
 Db 294 HIAKSKGNRYIVOMLLPVTSTRRL-RSSGVSPLHLAERNHDEVLALLSARFVNTPLA 352
 QY 298 -----INGQDKTALYNAVEKGNATWMDLIQCNPDIETCKDGEPLIKATKRNIEVY 352
 Db 353 PERARLYEDRRSSALYFVANNVYATELLQHGADP---NRDVISPLVAIRHGCLRTM 409
 QY 353 ELLDKGAKVSA 364
 Db 410 QLLDHGANIDA 421

RESULT 10
 TNRK2_HUMAN STANDARD: PRT; 1166 AA.
 AC 09H2K2: 09HAS4: 09H8F2:
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Tankyrase 2 (EC 2.4.2.30) (TANK2) (Tankyrase II) (TRF1-
 DE interacting ankyrin-related ADP-ribose polymerase 2) (Tankyrase-1like
 DE protein) (Tankyrase-related protein).
 OS TANK2 OR TNRK1 OR TANK2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP TISSUE: Fetal brain;
 RX MEDLINE=21072850; PubMed=11205898;
 RA Monz D., Munnia A., Comtesse N., Fischer U., Stendel W.-I., Feiden W.,
 RA Glasz B., Meese E.U.;
 RT "Novel tankyrase-related gene detected with meningioma-specific
 RT sera.";
 RL Clin. Cancer Res. 7:113-119(2001).
 RN [2]
 RP TISSUE: Breast carcinoma;
 RX MEDLINE=21190090; PubMed=11294570;
 RA Kurov A.N., Kuprash D.V., Petrov V.N., Vdovichenko K.K.,
 RA Scalan M.J., Jongeneel C.V., Lagarkova M.A., Nedospasov S.A.;
 RT "Cloning and characterization of TNRK1, a member of tankyrase gene
 RT family.";
 RL Genes Immun. 2:52-55(2001).
 RN [3]
 RP TISSUE: Liver;
 RX MEDLINE=21264473; PubMed=11278563;
 RA Lyons R.J., Deane R., Lynch D.K., Ye Z.-S.J., Sanderson G.M.,
 RA Eyre H.J., Sutherland G.R., Daly R.J.;
 RT "Identification of a novel human tankyrase through its interaction
 RT with the adaptor protein Gb14.";
 RL J. Biol. Chem. 276:17172-17180(2001).
 RN [4]
 RP SEQUENCE FROM N.A., AND SUBCELLULAR LOCALIZATION.

RC TISSUE=Placenta;
 RX MEDLINE=21443728; PubMed=11454873;
 RA Kaminker P.G., Kim S.-H., Taylor R.D., Zebardjian Y., Funk W.D.,
 RA Morin G.B., Yaswen P., Campisi J.;
 RT "TANK2, a new TRF1-associated poly(ADP-ribose) polymerase, causes
 RT rapid induction of cell death upon overexpression.";
 RL J. Biol. Chem. 276:35891-35899(2001).
 RN [5]
 RP SEQUENCE FROM N.A., FUNCTION, AND INTERACTION WITH TRF1 AND
 RP LNEP/OTASE.
 RC TISSUE=Skeletal muscle;
 RX MEDLINE=21661461; PubMed=11802774;
 RA Shodlo J.I., Lodish H.F., Chi N.-W.;
 RT "Tankyrase-2 oligomerizes with tankyrase-1 and binds to both TRF1
 RT (telomere-repeat-binding factor 1) and IRAP (insulin-responsive
 RT antipeptidase).";
 RL Biochem. J. 361:451-459(2002).
 RN [6]
 RP SEQUENCE FROM N.A.
 RA Yin Y., Gelmann E.P.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE FROM N.A.
 RA Tracey A.;
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 RN [8]
 RP SEQUENCE OF 838-1151 FROM N.A.
 RC TISSUE=Placenta;
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Takahashi-Fuji A., Hara H.,
 RA Tanase T., Nomura Y., Toshiya S., Komai F., Hara R., Takeuchi K.,
 RA Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,
 RA Wakamatsu A., Nakamura Y., Nagahara K., Masuno Y., Oshima A.;
 RT "NPO human cDNA sequencing project.";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 RN [9]
 RP FUNCTION, AND ADP-RIBOSYLATION.
 RX MEDLINE=21602874; PubMed=11739745;
 RA Cook B.D., Dynek J.N., Chang W., Shostak G., Smith S.;
 RT "Role for the related poly(ADP-ribose) polymerases tankyrase 1 and 2
 RT at human telomeres.";
 RL Mol. Cell. Biol. 22:332-342(2002).
 CC -1- FUNCTION: May regulate vesicle trafficking and modulate the
 CC subcellular distribution of SLC2A4/GLUT4-vesicles. Has PARP
 CC activity and can modify TRF1, and thereby contribute to the
 CC regulation of telomere length.
 CC -1- CATALYTIC ACTIVITY: NAD(+) + (ADP-D-riboseyl)(N)-acceptor =
 CC nicotinamide + (ADP-D-riboseyl)(N+1)-acceptor.
 CC -1- SUBUNIT: Oligomerizes and associates with TANKS. Interacts with the
 CC cytoplasmic domain of LNEP/OTase in SLC2A4/GLUT4-vesicles. Binds
 CC to the N-terminus of Gb14 and TRF1 with its ankyrin repeat
 CC region.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic, associated with the Golgi and
 CC with juxtanuclear SLC2A4/GLUT4-vesicles. Also found around the
 CC pericentriolar matrix of mitotic centrioles. During interphase, a
 CC small fraction of TANK2 is found in the nucleus, associated with
 CC TRF1.
 CC -1- TISSUE SPECIFICITY: Highly expressed in placenta, skeletal muscle,
 CC liver, brain, kidney, heart, thymus, spinal cord, lung, peripheral
 CC blood leukocytes, pancreas, lymph nodes, spleen, prostate, testis,
 CC ovary, small intestine, colon, mammary gland, breast and breast
 CC carcinoma, and in common-type meningioma. Highly expressed in
 CC fetal liver, heart and brain.
 CC -1- PTM: ADP-ribosylated (-auto).
 CC -1- SIMILARITY: BELONGS TO THE PARP FAMILY.
 CC -1- SIMILARITY: CONTAINS 15 ANK REPEATS.
 CC -1- SIMILARITY: CONTAINS 1 SAM DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

DR EMBL: AF305081; AAC35674.1; ALT_INT.
DR EMBL: AF264912; AAC44694.1; -
DR EMBL: AF329696; AAK13463.1; -
DR EMBL: AF342982; AAK25811.1; -
DR EMBL: AF309033; AAK82330.1; -
DR EMBL: AF348201; AAL40795.1; -
DR EMBL: AL359707; CAC78760.1; -
DR EMBL: AK023746; BAB1465.1; ALT_INT.
DR HSP; Q00420; IAWC.
DR Genew; HGNC:15677; TNKS2.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR001660; SAM.
DR Pfam; PF000023; ank; 20.
DR Pfam; PF00536; SAM; 1.
DR PRINTS; PR01415; ANKRYN.
DR SMART; SM00248; ANK; 15.
DR SMART; SM00454; SAM; 1.
DR PROSITE; PSS0088; ANK_REPEAT; 15.
DR PROSITE; PSS0297; ANK_REPEAT; 1.
DR PROSITE; PSS0105; SAM_DOMAIN; 1.
DR Transferrase; Glycosyltransferase; NND; Golgi stack; Telomere;
KW Nuclear protein; Repeat; ANK repeat; ADP-ribosylation.
DR repeat; 57.
DR repeat; 90.

FT	REPEAT	90	122	ANK 1.
FT	REPEAT	123	155	ANK 2.
FT	REPEAT	210	242	ANK 3.
FT	REPEAT	243	275	ANK 4.
FT	REPEAT	276	308	ANK 5.
FT	REPEAT	363	398	ANK 6.
FT	REPEAT	399	431	ANK 7.
FT	REPEAT	432	464	ANK 8.
FT	REPEAT	525	557	ANK 9.
FT	REPEAT	558	590	ANK 10.
FT	REPEAT	591	623	ANK 11.
FT	REPEAT	678	710	ANK 12.
FT	REPEAT	711	743	ANK 13.
FT	REPEAT	744	776	ANK 14.
FT	DOMAIN	873	936	ANK 15.
FT	DOMAIN	1023	1162	PARP.
FT	CONFLICT	331	337	KGSLQ -> ORPLVAA (IN REF. 1).
FT	CONFLICT	357	361	NFHRP -> IQAS (IN REF. 1).
FT	CONFLICT	966	966	Q -> P (IN REF. 8).
SO	SEQUENCE	1166	126917	MM; 4C8BB3B8D97CEFT04 CRC64;

Query Match	4.3%	Score 380.5;	DB 1;	Length 1166;
Best Local Similarity	24.0%;	Pred. No. 3.1e-13;		
Matches 157; Conservative	73;	Mismatches 206;	Indels 217;	Gaps 13;

[illegible]

```

Db      487  LGSNEADROLLLEAAKKAGDVEYTKLCTQVSVNCRIEEROSTPLHFAAGYKRVSVETLL 546
QY      225  KRNPNNWLDKDSNTMLMTASKEGHETIYODLLDAGTYVNIPIRSGDVTLLGAVRGSHVE 284
Db      547  QHGADYNAKKGGGLVPLHNACSGYGHAEVALLYKGAIVANADLMKFTPLHEAAAKKYE 606
QY      285  IVRALLOKYADIDIRQDNKFTALYUAVEKGNATWYRDILO----- 344
Db      607  ICKLLIQHGADPRPKKNRDGNTPRL-DLYKDGSDTDIDLLRQDALDLAAKGCGLARYKLL 664
QY      325  CNPDTETCPRDGE-----TPLIKATKRNRIEVEWELLDDGAYASAVDKKGDTPHL----- 374
Db      665  SSPDNVNC-RDTQGRHSTPLHLAAGYNINLEAEVYLLQGAIVANADQGGGLPLHNAASYG 723
QY      375  -----VAIRGSRRLAEILLRNPKDGRLLYRPNKAGE 406
Db      724  HYDVAAALLIKYNACVNATDKMAFTPLHEAQAQGRFT-QLCALLAHAGDPTL-----KNEQEG 779
QY      407  TPYNICSHQKSLITQIFGARHLSPETPDGMLGVDYSSALLADHLSPTIMOP 459
Db      780  TPLDLYSADYVSALLTAAAMPSPALTSCKRPQVLYNKVYRPGCATALALSSGSSSP 832

```

ID	ANKH_CHRV1	STANDARD;	PRT;	323 AA.
AC	006527;			
DT	01-FEB-1995 (Rel. 31, Created)			
DT	01-FEB-1995 (Rel. 31, last sequence update)			
DT	16-OCT-2001 (Rel. 40, last annotation update)			
DE	Ankyrin homolog precursor.			
OS	Chromatium vinosum.			
CC	Bacteria; Proteobacteria; gamma subdivision; Chromatiaceae;			
CC	Allochromatium.			
OX	NCBI_TaxID=1049;			
EN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=D / ATCC 17899 / DSM 180;			
RC	MEDLINE=93300842; PubMed=8390993;			
RA	Dolata M.M., van Beunuen J.J., Ambler R.P., Meyer T.E.,			
RA	Cusanovich M.A.;			
RT	"Nucleotide sequence of the heme subunit of flavocytochrome c from			
RT	the purple phototrophic bacterium, Chromatium vinosum. A 2.6-kilobase			
RT	pair DNA fragment contains two multiheme cytochromes, a flavoprotein,			
RT	and a homolog of human ankyrin."			
RL	J. Biol. Chem. 268:14426-14431(1993).			
CC	-1- SIMILARITY: CONTAINS 8 ANK REPEATS.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation			
CC	at the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce ,			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
EMBL:	L13419; AAA23315.1; -.			
DR	HSSP; PSS273; 1BD8.			
DR	InterPro; IPR002110; ANK.			
DR	Pfam; PF00023; ank; 8.			
DR	PRINTS; PR01415; ANKYRIN.			
DR	SMART; SM00248; ANK; 6.			
DR	PROSITE; PSS0088; ANK_REPEAT; 7.			
DR	PROSITE; PSS0297; ANK_REPEAT_REGION; 1.			
KW	Signal; ANK repeat; Repeat.			
FT	SIGNAL	1	28	POTENTIAL.
FT	CHAIN	29	323	ANKYRIN HOMOLOG.
FT	REPEAT	31	60	ANK 1.
FT	REPEAT	64	93	ANK 2.
FT	REPEAT	97	126	ANK 3.
FT	REPEAT	130	159	ANK 4.
FT	REPEAT	163	192	ANK 5.
FT	REPEAT	196	225	ANK 6.
FT	REPEAT			

FT REPEAT 229 258 ANK 7.
 FT SEQUENCE 262 291 ANK 8.
 SO SEQUENCE 323 AA; 33453 MW; 3976D3A49BD07DF CRC64;

Query Match
 Best Local Similarity 36.1%; Score 378.5; DB 1; Length 323;
 Matches 97; Conservative 43; Mismatches 128; Indels 1; Gaps 1;

QY 24 LLEKCDVDERNCGOTPLMLAEGNVEIVKELKNGANCNEDDNNWTALISAKESKH 83
 DB 51 LLEQGAHVDPNPGFTALFAVEGDDLEVALLSRGADVNRRTVAGCTPLFEAAEGH 110
 QY 84 IHIVEELLSGASILEHRDMGWTALMACYKGRDVEELLISGANPSYGLQSYPTII 143
 DB 111 IGISALLERGAHVHRTSRGWDALMASRHGTTDMVEQLFEKADPKADREGRT-ALM 169
 QY 144 MAARGHADIVHLLONGAKVNCSDYGTTPVMAARKHGLECVKHLAMGADVDOEGAN 203
 DB 170 QAASKETGYPLLLIGGADLEARDQGATALLIADQOAGAVETIAGIGADLDVADL 229
 QY 204 SMALIVAVKGYTQSVKELKRNPNVNLTKDGNALMASKEGHEIYODLLDAGTYV 263
 DB 230 GSTALLIAGHGVAVVERLLANGADPNRODRGCTTALMEAVTDAELIDRLIAGART 289
 QY 264 NIDPRSGDTVLIGAVGHEIVRALLQK 292
 DB 290 DLKDDADRTADIAVQGLGHQAIARLAPK 318

RESULT 12
 ANR3_HUMAN STANDARD; PRT; 832 AA.
 AC P57078: Q96KH0;
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 41, Last annotation update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Serine/threonine-protein kinase ANKRD3 (EC 2.7.1.-) (Ankyrin repeat domain protein 3) (PKC-delta-interacting protein kinase).
 GN ANKRD3 OR DIK.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE=Fetal kidney, and Fetal lung;
 RA Shimizu N., Kudoh J., Shibuya K.;
 RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.

RP SEQUENCE FROM N.A.
 RX MEDLINE=20289799; Pubmed=10830953;
 RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
 RA Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Soeda E.,
 RA Onki M., Takagi T., Sakaki Y., Tauden S., Blechschmidt K., Polley A.,
 RA Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
 RA Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
 RA Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
 RA Shtutman A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
 RA Minoshima S., Shimizu N., Nordsiek G., Hornischer K., Brandt P.,
 RA Scharte M., Schoen O., Desario A., Reichelt J., Kauer G., Bloeker H.,
 RA Ramser J., Beck A., Klages S., Hennig S., Risselmann L., Dagnod E.,
 RA Wehrach H., Reinhardt R., Gardiner K., Nizetic D., Francis F.,
 RA Lehman H., Reinhardt R., Yaspo M.-L.;
 RT "The DNA sequence of human chromosome 21.";
 RL Nature 405:311-319(2000).
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are
 CC produced by alternative splicing.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC -1- SIMILARITY: CONTAINS 10 ANK REPEATS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcements/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: AB047783; BAB56136.1; -;
 CC EMBL: AP001743; BAA95526.1; -;
 CC HSSP: P25963; 1IKK.
 CC Genew: HGNC:496; ANKRD3.
 DR MIM: 605705; -;
 DR InterPro: IPR002110; ANK.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR004040; STY_pkinase.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR Pfam: PF00023; ank; 10.
 DR Pfam: PF00069; pkinase; 1.
 DR PRINTS: PR00109; TYRKINASE.
 DR ProDom: PD000001; Euk_pkinase; 1.
 DR SMART: SM00248; ANK; 10.
 DR SMART: SM00221; STYK; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
 DR PROSITE: PS50088; ANK_REPEAT; 9.
 KW Transferase: Serine/threonine-protein kinase; ATP-binding; Repeat;
 KW ANK repeat; Alternative splicing.
 FT DOMAIN 22 286
 FT REPEAT 485 514 ANK 1.
 FT REPEAT 518 547 ANK 2.
 FT REPEAT 551 580 ANK 3.
 FT REPEAT 584 613 ANK 4.
 FT REPEAT 617 647 ANK 5.
 FT REPEAT 651 680 ANK 6.
 FT REPEAT 684 713 ANK 7.
 FT REPEAT 717 746 ANK 8.
 FT REPEAT 750 780 ANK 9.
 FT REPEAT 782 811 ANK 10.
 FT NP_BIND 28 36 ATP (BY SIMILARITY).
 FT BINDING 51 51 ATP (BY SIMILARITY).
 FT ACT_SITE 143 143 BY SIMILARITY.
 FT VARSPIC 278 325 MISSING (IN ISOFORM 2).
 FT CONFLICT 714 714 M -> V (IN REF. 1).
 SO SEQUENCE 832 AA; 91610 MW; 5D8FFD5F047FECB CRC64;

Query Match
 Best Local Similarity 4.1%; Score 360; DB 1; Length 832;
 Matches 109; Conservative 60; Mismatches 126; Indels 102; Gaps 6;

QY 29 KVDENRNECGOTPLMLAEGNVEIVKELKNGANCN----- 66
 DB 477 QVVDALDSGASLHVADEAGOECAKMLNNANPNLSNRGSLPLMAVRRYRGVYE 536
 QY 67 -----BDLNNWTALISAKESGHIHVEELLSGASILEHRDMGWTALMACYK 115
 DB 537 LILARKISVANADEDQWTALHFAAQNGESSRLLEKNASVNEVDFGRPMHYAOCQG 596
 QY 116 RTDVEELLISGANPSYGLQSYPTIIMAGRGADIVHLL-ONGAKVNCSDYGTTP 174
 DB 597 QENIVRILRRGVDSVLOG-KDAMLP.LHVAAGHLPYKLLAKQPGSVANAQTLDGRT 655
 QY 175 LVMARKHGLECVKHLAMGADVDOEGANSMTALIVAVKGYTQSVKELKRNPNVNLTD 234
 DB 656 LHLAQRGHRVARRLIIDCSVNCSLAQPLHVAEETGHTSTARLLIRGACKEMT 715
 QY 235 KDGNTALMASKEGHEIYQDLDAGTYVNIIDPRSGDTVLIGAVGHEIVRALLQK 294
 DB 716 SDGYTALHLAA-----RNGHLATVKLLVEEKA 742
 QY 295 DIDIRGDKNTALIVAVEKGNATWRTDIQCNPTIECTKQDEPLIKATKRNIEVEL 354
 DB 743 DVLAIRPQLMOTALHLAAAGHSEVVEIV-----SADVIDL 778

OY 355 LIDKAKASAVDKDPTPLHVALRGSRRLAELLRN 391
Db 779 F-----DEQGLSALHLAAQGRHAQVETLLKH 805

RESULT 13
RNSL_HUMAN STANDARD; PRT; 741 AA.
AC 005823;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 2-5A-dependent ribonuclease (EC 3.1.26.-) (2-5A-dependent RNase)
DE (Ribonuclease L) (RNase L) (Ribonuclease 4).
GN RNSAEL OR RNSA.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND MUTAGENESIS OF LYS-240 AND LYS-274.
RC TISSUE=Kidney;
RC MEDLINE=93201598; PubMed=7680958;
RA Zhou A., Hassel B.A., Silverman R.H.;
RT "Expression cloning of 2-5A-dependent RNase: a uniquely regulated
RT mediator of interferon action.";
RL Cell 72:753-765(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RC MEDLINE=20515260; PubMed=11063255;
RA Zhou A., Nie H., Silverman R.H.;
RT "Analysis and origins of the human and mouse RNase L genes: mediators
RT of interferon action.";
RL Mamm. Genome 11:989-992(2000).
RN [3]
RP CHARACTERIZATION OF RNSAEL ACTIVITY.
RC MEDLINE=94245737; PubMed=7514601;
RA Dong B., Xu L., Zhou A., Hassel B.A., Lee X., Torrence P.F.,
RA Silverman R.H.;
RT "Intrinsic molecular activities of the interferon-induced 2-5A-
RT dependent RNase.";
RL J. Biol. Chem. 269:14153-14158(1994).
RN [4]
RP FUNCTION, AND SUBCELLULAR LOCATION.
RC PubMed=11585831;
RA Le Roy F., Bisbal C., Silhol M., Martinand C., Leblou B.,
RA Salehzada T.;
RT "The 2-5A/RNase L/RNase L inhibitor (RNI) pathway regulates
RT mitochondrial mRNA stability in interferon alpha-treated H9 cells.";
RL J. Biol. Chem. 276:48473-48482(2001).
RN [5]
RP ERRATUM.
RA Le Roy F., Bisbal C., Silhol M., Martinand C., Leblou B.,
RA Salehzada T.;
RL J. Biol. Chem. 277:13354-13354(2002).
RN [6]
RP REVIEW.
RC PubMed=9856285;
RA Castell J., Wood K.A., Youle R.J.;
RT "The 2-5A system in viral infection and apoptosis.";
RL Biomed. Pharmacother. 52:386-390(1998).
RN [7]
RP MUTAGENESIS OF LYS-392.
RC PubMed=9862963;
RA Dong B., Silverman R.H.;
RT "Alternative function of a protein kinase homology domain in 2',
RT 5'-oligoadenylate dependent RNase L.";
RL Nucleic Acids Res. 27:439-445(1999).
RN [8]
RP MUTAGENESIS OF HIS-583; PRO-584; TRP-632; ASP-661; ARG-667 AND
RP HIS-672.

RX PubMed=11333017;
RA Dong B., Niwa M., Walter P., Silverman R.H.;
RT "Basis for regulated RNA cleavage by functional analysis of RNase L
RT and Ire1p.";
RL RNA 7:361-373(2001).
RN [9]
RP VARIANTS SER-59; PHE-406; GLN-462 AND GLU-541.
RX PubMed=11941539;
RA Roekman A., Ikonen T., Seppälä E.H., Nupponen N., Autio V.,
RA Mononen N., Bailey-Wilson J., Trent J., Carpen J., Matikainen M.P.,
RA Koivisto P.A., Tammela T.L.J., Kallioniemi O.-P., Schleutker J.;
RT "Germline alterations of the RNSAEL gene, a candidate HPC1 gene at
RT 1q25, in patients and families with prostate cancer.";
RL Am. J. Hum. Genet. 70:1299-1304(2002).
CC -1- FUNCTION: Endoribonuclease, mediator of interferon action, which
CC play a role in mediating resistance to virus infection and
CC apoptosis. Might play a central role in the regulation of mRNA
CC turnover.
CC -1- CATALYTIC ACTIVITY: Cleaves 3' of UPNp dimers, with preference for
CC U and UA sequences, to sets of discrete products ranging from
CC between 4 and 22 nucleotides in length.
CC -1- COTRFACTOR: Optimal RNA cleavage rates requires the presence of
CC either manganese or magnesium and ATP.
CC -1- ENZYME REGULATION: After binding to 2-5A (5'-phosphorylated 2',5'-
CC linked oligoadenylates) the homodimerization and subsequent
CC activation occurs. Inhibited by RNase L inhibitor.
CC -1- SUBUNIT: Monomer (inactive form) or homodimer.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic and mitochondrial.
CC -1- TISSUE SPECIFICITY: Highly expressed in spleen and thymus followed
CC by prostate, testis, uterus, small intestine, colon and peripheral
CC blood leukocytes.
CC -1- INDUCTION: By interferons.
CC -1- DOMAIN: The kinase domain allows the homodimerization.
CC constitute the 2-5A binding domain.
CC -1- DISEASE: The nine ankryrin repeats also called 2-5A sensor
CC susceptibility for HPC1.
CC -1- DISEASE: Germline defects in RNSAEL are linked to cancer-
CC susceptibility for HPC1.
CC -1- SIMILARITY: CONTAINS 9 ANK REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L10381; AAA18032.1; -;
DR HSSP: P80144; ZMYO.
DR Genew: HGNC:10050; RNSAEL.
DR MIM: 180435; -;
DR InterPro: IPR002110; ANK.
DR InterPro: IPR000719; Euk_Pkinase.
DR Pfam: PF00023; ank; 8.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Euk_pkinase; 1.
DR SMART: SM00248; ANK; 8.
DR PROSITE: PS50088; ANK_REPEAT; 6.
DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
KW Hydrolase; Nuclease; RNA-binding; Endonuclease; Zinc-finger; Repeat;
KW ANK repeat; Polymorphism.
FT REPEAT 24
FT REPEAT 58 87 ANK 1.
FT REPEAT 91 120 ANK 2.
FT REPEAT 124 153 ANK 3.
FT REPEAT 167 197 ANK 4.
FT REPEAT 201 234 ANK 5.
FT REPEAT 238 268 ANK 6.
FT REPEAT 272 301 ANK 7.
FT REPEAT 303 329 ANK 8.
FT REPEAT 365 586 ANK 9.
FT DOMAIN 587 741 PROTEIN KINASE LIKE.
FT RIBONUCLEASE.

```

FT DOMAIN 229 242 2-5A BINDING (P-LOOP) 1.
FT DOMAIN 253 275 2-5A BINDING (P-LOOP) 2.
FT ZN_FING 395 444 C6-TYPE (ATYPICAL).
FT VARIANT 59 59 G-> S.
FT VARIANT 406 406 /FID-VAR_013509.
FT VARIANT 406 406 S-> F.
FT VARIANT 462 462 /FID-VAR_013510.
FT VARIANT 462 462 R-> Q (RISK FACTOR FOR HPC1; IN
FT DBSNP:4869077).
FT VARIANT 541 541 /FID-VAR_012056.
FT VARIANT 541 541 D-> E (IN DBSNP:627928).
FT MUTAGEN 240 240 /FID-VAR_012057.
FT MUTAGEN 240 240 K->N: REDUCED 2-5A BINDING ACTIVITY;
FT MUTAGEN 274 274 K->N: REDUCED 2-5A BINDING ACTIVITY;
FT MUTAGEN 274 274 K->N: REDUCED 2-5A BINDING ACTIVITY;
FT MUTAGEN 392 392 K->N: REDUCED 2-5A BINDING ACTIVITY;
FT MUTAGEN 392 392 K->R: COMPLETE LOSS OF ENZYMACTIC ACTIVITY
AND ENZYME DIMERIZATION. NO CHANGE IN
FT BINDING TO 2-5A AND RNA.
FT MUTAGEN 583 583 H->A: NO CHANGE IN ENZYMACTIC ACTIVITY.
FT MUTAGEN 584 584 P->A: NO CHANGE IN ENZYMACTIC ACTIVITY.
FT MUTAGEN 632 632 W->A: NO CHANGE IN ENZYMACTIC ACTIVITY.
FT MUTAGEN 661 661 D->A: COMPLETE LOSS OF ENZYMACTIC
ACTIVITY.
FT MUTAGEN 667 667 R->A: COMPLETE LOSS OF ENZYMACTIC
ACTIVITY. NO CHANGE IN 2-5A BINDING AND
FT ENZYME DIMERIZATION.
FT MUTAGEN 672 672 H->A: COMPLETE LOSS OF ENZYMACTIC
ACTIVITY. NO CHANGE IN 2-5A BINDING
FT ACTIVITY AND ENZYME DIMERIZATION.
FT SEQUENCE 741 AA; 83532 MW; 91385EA307E3CELD CRC64;
Query Match 4.0%; Score 357.5; DB 1; Length 741;
Best Local Similarity 27.4%; Pred. No. 2.8e-12;
Matches 121; Conservative 64; Mismatches 128; Indels 129; Gaps 13;
QY 42 LMLAEGNVEIKKELKNCNLEEDLWTLALISKRGHHIYELLKSGASLEHRD 101
DB 29 LIRKAVQEDVDLVQQLLEGANVFOE-----E 56
QY 102 MGGWTALMAYCYKRTDVVELLSHGANPSVTGLQSYPIIWAAGRADIYHLILONG 161
DB 57 EGGWTPILHNAVQMRREDIVELLRHGADPYL----- 87
QY 162 AKVNCSDRYGTPPLVMAARKGHLECYVHLLAMGADVDEGANSMTALIVAKGYIOSVK 221
DB 88 -----RRKNGATPFIILAIAGSVLRLFLFSKGDVNECFYGTAFMEAAVYGVKVALK 142
QY 222 EILKRNPNVNL-----TRKDGNTALMIAKSGEHIYVODLD-AGTYVNIPIRSG 270
DB 143 FLIRKAGANVLRRTKEDQERLRKGGATLMDAAEKHVEVLKLLIDEMGADVNAACNMNG 202
QY 271 DTVLIGAVRG--GHVE-IYRALLQKYADIDIRGQDKTALYAWAVERGNATWVRDILQCN 326
DB 203 RNALILHALSSDSDVFAITFLILDHGADVNERGKPTLLAVEKKHGLIVQRLLEOE 262
QY 327 -----PDREICRKDEPILIKATKRNINEVVELLDKAKAYSAVKKDPTPLHVALIRRSR 382
DB 263 HIETINDV--SDKRTALLAVELLKKRIKIELCKRASTDC-----GDLVA-TARRNYDH 314
QY 383 RLAEELLNRP-----KDRGLYVPRNAGEFPYIIDSOSKSI 419
DB 315 SLIVAVLLSHGAKEDFHPAPEDWKRPOSSHWGALALDLIRIYRP-MIGLTKPFIDIKYR--- 370
QY 420 LTOIFGARHLSPLETGDMIGY 441
DB 371 -----IADTSEGGIYIGF 383

```

```

ID RN5A_MOUSE STANDARD: PRT: 735 AA.
AC 005921; OPERU;
DT 01-NOV-1995 (rel. 32, Created)
DT 15-JUN-2002 (rel. 41, Last sequence update)
DT 15-JUN-2002 (rel. 41, Last annotation update)
DE 2-5A-dependent ribonuclease (EC 3.1.26.-) (2-5A-dependent RNase)
DE (Ribonuclease L) (RNase L) (Ribonuclease 4).
GN RNASEL OR RN5A.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=C3H/An; TISSUE=Adipose tissue;
RX MEDLINE=20515260; PUBMED=11063255;
RA Zhou A., Nie H., Silverman R.H.;
RT "Analysis and origins of the human and mouse RNase L genes: mediators
of interferon action.";
RL Mamm. Genome 11:989-992(2000).
[2]
RN SEQUENCE OF 1-679 FROM N.A.
RP MEDLINE=93201598; PUBMED=7680958;
RX Zhou A., Hassel B.A., Silverman R.H.;
RT "Expression cloning of 2-5A-dependent RNase: a uniquely regulated
mediator of interferon action.";
RL Cell 72:753-765(1993).
[3]
RN FUNCTION, AND SUBCELLULAR LOCATION.
RP PUBMED=11585831;
RX Le Roy F., Bisbal C., Silhol M., Martinand C., Lebou B.,
RA Salehzada T.;
RT "The 2-5A/RNase L/RNase L inhibitor (RNI) pathway regulates
mitochondrial mRNA stability in interferon alpha-treated H9 cells.";
RL J. Biol. Chem. 276:48473-48482(2001).
[4]
RN ERROR.
RA Le Roy F., Bisbal C., Silhol M., Martinand C., Lebou B.,
RA Salehzada T.;
RL J. Biol. Chem. 277:13354-13354(2002).
CC 1- FUNCTION: Endoribonuclease, mediator of interferon action, which
CC play a role in mediating resistance to virus infection and
CC apoptosis. Might play a central role in the regulation of mRNA
CC turnover.
CC 1- CATALYTIC ACTIVITY: Cleaves 3' of UPNP dimers, with preference for
CC UU and UA sequences, to sets of discrete products ranging from
CC between 4 and 22 nucleotides in length.
CC 1- COFACTOR: Optimal RNA cleavage rates requires the presence of
CC either manganese or magnesium and ATP.
CC 1- ENZYME REGULATION: After binding to 2-5A (5'-phosphorylated 2',5'-
CC linked oligoadenylates) the homodimerization and subsequent
CC activation occurs. Inhibited by RNase L inhibitor.
CC 1- SUBUNIT: Monomer (inactive form) or homodimer.
CC 1- SUBCELLULAR LOCATION: Cytoplasmic and mitochondrial.
CC 1- TISSUE SPECIFICITY: Expressed in spleen, thymus, lung, testis,
CC kidney, liver and heart.
CC 1- INDUCTION: By interferons.
CC 1- DOMAIN: A single active nuclease domain in a dimer is
CC sufficient for ribonuclease activity.
CC 1- DOMAIN: The kinase domain allows the homodimerization.
CC 1- DOMAIN: The nine ankyrin repeats also called 2-5A sensor
CC constitute the 2-5A binding domain.
CC 1- SIMILARITY: CONTAINS 9 ANK REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF281045; MG33708.1; -

```

DR EMBL: L10382; AAA37117.1; -
 DR HSSP: P42773; 11HB.
 DR MGI:1098272; Rnasel.
 DR InterPro: IPR002110; Ank.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR Pfam: PF00023; ank. 8.
 DR Pfam: PF00069; pkinase. 1.
 DR PRINTS: PRO1415; ANKYRIN.
 DR SMART: SM00248; ANK. 8.
 DR SMART: SM00220; S_TKC. 1.
 DR SMART: SM00219; Tyrc. 1.
 DR PROSITE: PS50088; ANK_REPEAT. 7.
 DR PROSITE: PS50297; ANK_REPEAT_REGION. 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM. 1.
 DR Hydrolase; Nuclease; RNA-binding; Endonuclease; Zinc-finger;
 KW Repeat; ANK repeat.
 FT REPEAT 24 ANK 1.
 FT REPEAT 58 ANK 2.
 FT REPEAT 91 ANK 3.
 FT REPEAT 124 ANK 4.
 FT REPEAT 167 ANK 5.
 FT REPEAT 201 ANK 6.
 FT REPEAT 238 ANK 7.
 FT REPEAT 272 ANK 8.
 FT REPEAT 303 ANK 9.
 FT REPEAT 364 ANK 9.
 FT DOMAIN 584 PROTEIN KINASE LIKE.
 FT DOMAIN 585 RIBONUCLEASE.
 FT DOMAIN 229 2-5A BINDING (P-LOOP) 1.
 FT DOMAIN 253 2-5A BINDING (P-LOOP) 2.
 FT ZN_FING 401 C6-TYPE (POTENTIAL).
 FT SEQUENCE 735 AA; 83274 MW; B6632F4A5B50F711 CRC64;

Query Match 4.0%; Score 356; DB 1; Length 735;
 Best Local Similarity 32.2%; Pred. No. 3,4e-12;
 Matches 110; Conservative 57; Mismatches 111; Indels 64; Gaps 11;

QY 73 TALISASKEGHIHYEELKSGALSIEH-RDMGQGTALMMACYKRTYVVELLSHGANS 131
 DB 27 SSLIKAAQGVVAVRQQLLEKGAJANACEDTWTGTPHNAVAQAGRVIVNLLSHGADP 85
 QY 132 VTGLQYSYPIYMAAGRHADIVHLLONGAKVNCSDKYGTPTPLVMAARKGHLECYVHL 191
 DB 86 -----HRRKKNGA-----TPITAGIGDVAKLEILL 112
 QY 192 AMGADVDOEGANSMTALIVAVKGYTQSVKEILKRNPNVL---TDKD-----GNTAL 241
 DB 113 SCGADVNECEDNGFTAFMEAAERGNABALRPLFAKGANVNLRRQTTDKRRKLGSGATL 172
 QY 242 MIAKEGHEIEVOLL-DAGTYVNIIPRSGDTVLIGAVRGHV---EIVRALLQYADI 296
 DB 173 MSAAEKGLHEVLRILLDMKRAEVDARNMGRNALIRTLWDCENVEITSIILIOHADV 232
 QY 297 DIRGQDKTALYVAVKGNATVVDIIQ---CNPDTICTGDKGTEPLIKATKMRNEVVE 353
 DB 233 NVREGRKTPLIANAVERKHGIVOMLSREGINIDAR--DNEGKTALLIADVOKOLEIYO 290
 QY 354 LLLDKGAKVSAVDKGDPTLHVALIRGSRRLAEILL---RNP 392
 DB 291 LLEKGA-----DKCDLWVIARNDHYLVKLLPYANP 326

RESULT 15
 ANR6_HUMAN STANDARD; PRT; 692 AA.
 AC Q9Y2G4; Q9Y2G4; Q9Y2G4; Q9Y2G4;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Ankyrin repeat domain protein 6.
 GN ANKRD6 OR KIAA0957.
 OS Homo sapiens (Human).

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 CC NCBI_TaxID=9606;
 CC [1]
 CC SEQUENCE FROM N.A. (ISOFORM 1).
 CC TISSUE=Brain;
 CC MEDLINE=99246063; PubMed=10231032;
 CC Nagase T., Ishikawa K.-I., Suyama K., Kikuno R., Hirose S.,
 CC Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
 CC "Prediction of the coding sequences of unidentified human genes. XIII.
 CC The complete sequences of 100 new cDNA clones from brain which code
 CC for large proteins in vitro.";
 CC DNA Res. 6:63-70(1999).
 CC [2]
 CC SEQUENCE OF 87-692 FROM N.A. (ISOFORM 2).
 CC TISSUE=Testis;
 CC Poustka A., Klein M., Mewes H.-W., Gassenhuber J., Wiemann S.;
 CC Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 CC [3]
 CC SEQUENCE OF 140-692 FROM N.A. (ISOFORM 1).
 CC Tracey A.;
 CC Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; MAY BE
 CC -1- PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- SIMILARITY: CONTAINS 8 ANK REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----

DR EMBL: AB023174; BAA76801.1; -
 DR EMBL: AL117504; CAB55968.1; -
 DR EMBL: AL096678; CAB86658.1; -
 DR HSSP: Q00420; IAWC.
 DR GeneW: HGNC:17280; ANKRD6.
 DR InterPro: IPR002110; ANK.
 DR Pfam: PF00023; ank. 8.
 DR PRINTS: PRO1415; ANKYRIN.
 DR SMART: SM00248; ANK. 6.
 DR PROSITE: PS50088; ANK_REPEAT. 6.
 DR PROSITE: PS50297; ANK_REPEAT_REGION. 1.
 KW Repeat; ANK repeat; Alternative splicing.
 FT REPEAT 9 ANK 1.
 FT REPEAT 41 ANK 2.
 FT REPEAT 74 ANK 3.
 FT REPEAT 107 ANK 4.
 FT REPEAT 140 ANK 5.
 FT REPEAT 173 ANK 6.
 FT REPEAT 206 ANK 7.
 FT REPEAT 239 ANK 8.
 FT REPEAT 264 ANK 8.
 FT VARSPLIC 264
 FT Q -> QYLFSEKSGSLRKKRRLKEERRAQSVPDEVAQSK
 FT (IN ISOFORM 2).
 FT N -> I (IN REF. 2).
 FT CONFLICT 87
 FT SEQUENCE 692 AA; 75722 MW; D32115487A80839F CRC64;

Query Match 3.9%; Score 349.5; DB 1; Length 692;
 Best Local Similarity 29.2%; Pred. No. 7e-12;
 Matches 120; Conservative 67; Mismatches 177; Indels 47; Gaps 11;

QY 95 ASLEHRDMGQWTALMMACYKRTYVVELLSHGANSVTGLQYSYPIYMAAGRHADIV 154
 DB 8 AALSER-----LLVAYKQGTENVYQILNKGARVAVT--KHRTPLHLAANKGHLPIV 58
 QY 155 HLLONGAKVNCSDKYGTPTPLVMAARKGHLECYVHLLAMGADVDOEGANSMTALIVAVK 214
 DB 59 QILAKAGCDLLDYQDDGDTALHRAVYGNTEILALIHGEGCALDRQDGDGTALHRSWH 118
 QY 215 GYTQSVKELKRNPNVNLTDGNTALMIAKEGHEIEVOLL-DAGTYVNIIPRSGDTVL 274

```

Db      119 GFSQSAKLIVYKAGANVLAANKKAGNFALHLAQCNSHSQSTRVLLLAGSRADLKNNAQDTCL 178
Oy      275 IGVAVRGHVEIYRALLQKTAADIDIGQDNKTAIYVAEKGATVRILOQCNPTLCTK 334
Db      179 HVAAIRNNHLSIIRLLTAFCSVHEKNOAGDADLHVAAALNKKYAKLILEAGADTTIVNN 238
Oy      335 DGEPLIKATKMMINIEVVELLDKCAKYSAYDKKQDF-LHVATRGSRRLAEILLNPK 393
Db      239 AGQPLLETARYHNHPYA-LLTIKAPQGSV--SAGDTPSSQVAAKREAREPLASPE 295
Oy      394 -----DGRLLYRNPKAG-----ETPYNIDCSHOKSITLITQIGARHLSPTEND----- 435
Db      296 PRAKDDRRKRSRKYSAFSDPPRPADDOOPGHOKNL-----HANHPPKRRNRHRCSSPPP 349
Oy      436 -GAMLDGYLSALADILSEPTMOPPICGYAOWGSGSKFLKLLKDEPK 485
Db      350 PHEFRAYOLT--LYNGDKGKWAQAPIN-----GCCEPLINKLENGOLE 391

```

Search completed: July 1, 2003, 14:35:02
Job time : 33 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2003, 14:34:05 ; Search time 43 Seconds
(without alignments)
3834.198 Million cell updates/sec

Title: US-10-021-571-2
Perfect score: 8853
Sequence: 1 MSVLISQSVINVEENIPAL.....ELHAASSSTGGEERESIL 1715

Scoring table: BLOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5073.5	57.3	1031	T43458	hypothetical prote
2	1848.5	20.9	1398	T21884	hypothetical prote
3	1024	11.6	246	T46392	hypothetical prote
4	555	6.3	1881	1 S4HUK	ankyrin 1, erythro
5	555	6.3	426	AE2149	hypothetical prote
6	553	6.2	1856	B35049	ankyrin 1, erythro
7	553	6.2	1880	A35049	ankyrin 1, erythro
8	552	6.2	1961	T42716	ankyrin 3, splice
9	551.5	6.2	1940	T42715	ankyrin 3, splice
10	551	6.2	1943	T42713	ankyrin 3, splice
11	550	6.2	3924	S37431	ankyrin 2, neurona
12	527.5	6.0	1862	1 A55575	ankyrin 3, long sp
13	526	5.9	4377	2 A55575	ankyrin 3, long sp
14	516.5	5.8	1848	2 S37771	ankyrin, erythrocy
15	510.5	5.8	1765	T42714	ankyrin 3, splice
16	479	5.4	2039	T15347	ankyrin-related un
17	464.5	5.2	1549	T13940	ankyrin - fruit fl
18	429	4.8	2584	T24158	hypothetical prote
19	429	4.8	2606	T24157	hypothetical prote
20	414.5	4.7	1423	1 T37275	death-associated p
21	409.5	4.6	791	T42691	hypothetical prote
22	397	4.5	1062	T30255	inversin - mouse
23	384.5	4.3	557	T46507	hypothetical prote
24	382	4.3	1062	T14151	inv protein - mous
25	378.5	4.3	323	2 B47169	ankyrin-like repea
26	361	4.1	397	2 T46445	hypothetical prote
27	356	4.0	679	2 B45771	2-5a-dependent RNA
28	347.5	3.9	741	2 A45771	2-5a-dependent RNA
29	338.5	3.8	991	2 T25412	hypothetical prote

30	335	3.8	1401	2 S11527	alpha-latrotoxin p
31	325	3.7	1188	2 T19552	hypothetical prote
32	323.5	3.6	441	2 S74626	erythroid ankyrin
33	321.5	3.6	934	1 H71274	probable ankyrin -
34	302	3.4	1058	2 D82654	ankyrin-like prote
35	301	3.4	2004	2 D88948	protein ZK1005.1 f
36	300.5	3.4	1411	2 S30355	alpha-latroinsecto
37	296.5	3.3	627	2 E86190	hypothetical prote
38	294	3.3	1184	2 T00253	gene Ankhzn protei
39	292	3.3	606	2 AC2508	hypothetical prote
40	288	3.3	1083	2 S48460	probable membrane
41	287	3.2	517	2 T48283	ankyrin-like prote
42	286.5	3.2	662	2 E84725	ankyrin-like prote
43	285	3.2	1001	2 S30385	G9a protein - huma
44	279.5	3.2	368	2 T18184	ankyrin repeat pro
45	279	3.2	543	2 C86212	hypothetical prote

ALIGNMENTS

RESULT 1
T43458
hypothetical protein DKFZP434F0621.1 - human
C:Species: Homo sapiens (man)
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jan-2000
C:Accession: T43458
R:Blum, H.; Bauersachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, December 1999
A:Reference number: 222517
A:Accession: T43458
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1031 <AAA>
A:Cross-references: EMBL:AL133620
A:Experimental source: adult testis; clone DKFZP434F0621
A:Genetics:
A>Note: DKFZP434F0621.1

Query Match	Similarity	Score	DB 2:	Length
Best Local Similarity	95.2%	Pred. No. 4.8e-293;		
Matches	974;	Conservative	13;	Indels 13; Gaps 1;
QY	1	MSVLISQSVINVEENIPAL	KALLKCKVDENEGGPTLMAAGNVEIKELKN	60
DB	1	MSVLISQSVINVEENIPAL	KALLKCKVDENEGGPTLMAAGNVEIKELKN	60
QY	61	GANCNLEDLDNMTALISAK	EGHIIYEEELKSGASLEHRDGMGTALMAACYKGRDYY	120
DB	61	GANCNLEDLDNMTALISAK	EGHIIYEEELKSGASLEHRDGMGTALMAACYKGRDYY	120
QY	121	ELLISHGANPSVTGLQTSY	PIIYAAGRGADYIHLHLONGAKYNSDKTPTPLVAAR	180
DB	121	ELLISHGANPSVTGLQTSY	PIIYAAGRGADYIHLHLONGAKYNSDKTPTPLVAAR	180
QY	181	KGHLECYVHLLAMGADV	DOEGANSMTALIVAVKGYTQSYKEILKRPNVNLDPKDGNTA	240
DB	181	KGHLECYVHLLAMGADV	DOEGANSMTALIVAVKGYTQSYKEILKRPNVNLDPKDGNTA	240
QY	241	LMISAKEGHIEIVODLDA	GTYYNIIPRSGDTVLIGAVRGHVEIVRALLOKYADIDIRG	300
DB	241	LMISAKEGHIEIVODLDA	GTYYNIIPRSGDTVLIGAVRGHVEIVRALLOKYADIDIRG	300
QY	301	QDNKTALYMAVEKGNAT	VRRIIDCNPDTEICTDGETPLIKAKMNIIEVELLDKGA	360
DB	301	QDNKTALYMAVEKGNAT	VRRIIDCNPDTEICTDGETPLIKAKMNIIEVELLDKGA	360
QY	361	KVSADVKKGDTPLVAIR	GRSRLAELLRNPKGRILYPRNKAGEPPYNDISHOKSIL	420
DB	361	KVSADVKKGDTPLVAIR	GRSRLAELLRNPKGRILYPRNKAGEPPYNDISHOKSIL	420
QY	421	TOIFGARHLSPETEDG	MDLGYDLYSSALADILSEPTMQPPICVGLYQWQSSGSFLLKTL	480
DB	421	TOIFGARHLSPETEDG	MDLGYDLYSSALADILSEPTMQPPICVGLYQWQSSGSFLLKTL	480

Db 421 TQIFGARHLSPETDGMGLDYSSALADILSEPTOMPICVGLYAQMSGKSPFLK 480
 QY 481 EDEMTFAGGOTEPLFQPSMLIVELTLICGSLGYFPPDTNLAISTSPALTYIF 540
 Db 481 EDEMTFAGGOTEPLFQPSMLIVELTLICGSLGYFPPDTNLAISTSPALTYIF 540
 QY 541 FVIVFGRRGEGESNMWMAVSTRLARHIGYDELLEFKLMFVNPPELPQOTKALPVRELF 600
 Db 541 FVIVFGRRGEGESNMWMAVSTRLARHIGYDELLEFKLMFVNPPELPQOTKALPVRELF 600
 QY 601 TYNRLSSVGGESTLAEMATISDCEREPGLATRLRVRTESGKKKKTKCCLPS 660
 Db 601 TYNRLSSVGGESTLAEMATISDCEREPGLATRLRVRTESGKKKKTKCCLPS 660
 QY 661 FVIFVYGCIIAGITLLAIFRVDPKHLTVANILISVYGLAVLNCRTMVOYDLSL 720
 Db 661 FVIFVYGCIIAGITLLAIFRVDPKHLTVANILISVYGLAVLNCRTMVOYDLSL 720
 QY 721 NSQRRRLHVAASKLHKLSEGFMYLKCCEVELMARMATIDFTONOTRLVYIIDGLDAC 780
 Db 721 NSQRRRLHVAASKLHKLSEGFMYLKCCEVELMARMATIDFTONOTRLVYIIDGLDAC 780
 QY 781 EODKYLQMLDTRVLFSGPFLAIFASDPHIIKAINONLSVLRDSINGHDIYRNIVH 840
 Db 781 EODKYLQMLDTRVLFSGPFLAIFASDPHIIKAINONLSVLRDSINGHDIYRNIVH 840
 QY 841 LPVFNLSRGLSNARKFLVTSATNGDITGSDTGTQEDTDRRYSONSLGEMTKLGSKALN 900
 Db 841 LPVFNLSRGLSNARKFLVTSATNGDITGSDTGTQEDTDRRYSONSLGEMTKLGSKALN 900
 QY 901 RRDYRRROMORTITRQMSFDITKLVTEDWFSDISPQTMRLNIVSVTGRLLRANOIT 960
 Db 901 RRDYRRROMORTITRQMSFDITKLVTEDWFSDISPQTMRLNIVSVTGRLLRANOIT 960
 QY 961 FMDWDLASMINTEOMPYRTSMLILYLETBGLPQOMLKTMYER-----IS 1007
 Db 961 FMDWDLASMINTEOMPYRTSMLILYLETBGLPQOMLKTMYER-----IS 1007
 QY 1008 KNI 1010
 Db 1021 KSV 1023
 Db 1021 KSV 1023
 RESULT 2
 T21884
 hypothetical protein F36H1.2 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T21884
 R:Kershaw, J.
 Submitted to the EMBL Data Library, January 1996
 Reference number: Z19482
 Accession: T21884
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1398 <WILL>
 A:Cross-References: EMBL:Z68760; PIDN:CAA92996.1; GSPDB:GNO0022; CESP:F36H1.2
 A:Experimental source: clone F36H1
 C:Genetics:
 A:Gene: CESP:F36H1.2
 A:Map position: 4
 A:Introns: 38/3; 60/3; 94/3; 152/3; 191/3; 257/2; 290/3; 323/3; 379/1; 439/1; 485/2; 552/2
 Query Match 20.9%; Score 1848.5; DB 2; Length 1398;
 Best Local Similarity 31.5%; Pred. No. 3.4e-101;
 Matches 461; Conservative 272; Mismatches 459; Indels 271; Gaps 37;
 QY 34 RNEGQTPMLAEEGNGEIVKELKANGCNLEDLN--WTALISASKEGHIIIVEEL 91
 Db 34 RNEGQTPMLAEEGNGEIVKELKANGCNLEDLN--WTALISASKEGHIIIVEEL 91
 QY 58 RNANESLITVAVRSGNTAVAKQLQDLPDA-IDETDEGMSALLNAAHGHVDIVRLII 116
 Db 58 RNANESLITVAVRSGNTAVAKQLQDLPDA-IDETDEGMSALLNAAHGHVDIVRLII 116
 QY 92 KSGASLEHRDGMGWALMWACYKGTVDVVELLSGANGPSVTGLQSYVPIITMAAGRHA 151
 Db 92 KSGASLEHRDGMGWALMWACYKGTVDVVELLSGANGPSVTGLQSYVPIITMAAGRHA 151

Db 117 DNGASVDQDPLMGSPMLMAVYKNNHLDVYDLVN----- 150
 QY 152 DIYHLLQNGAKNCSKYCTEPLVMAARGHIECYKHILLAMGADVQDGAANSMTALYA 211
 Db 151 -----AKVNACDKRESTALIMARKGHLPVQULLNSGAEDVAGMSTRALMA 200
 QY 212 VKGGYTSVKEILLRNPNVMTDRKDNALIMASKGHEIYVODLDAGTYVNPDRSG 271
 Db 201 TRGFIQVVELLTLRENVNAVADONGTLALGMAARGVADICSLNSGAFVNGCRPN 260
 QY 272 TVLIGAVRGHVEIVRALLOKVIDIRGDNKTALYAVEKGNATMYVDILOCNDTEI 331
 Db 261 WILTSVARSNGAAIVRMILDKFADINCODESEKTPHLIDKSFNDIAVILTEKKNNEL 320
 QY 332 CTGDEGEPILKATKRMIEVEVVELLDKGAVSADVKKGDPPLHATRGSRRLAEELN 391
 Db 321 KNGDETPPLRAKCRNVHICTYLMSEGAALAVNDCGNALALANSRRLTQALLSN 380
 QY 392 PKDGRLLYRNKAGETPPYNIIDCSHOKSILTQIFGARHLSPETDGM--LAGDYSSAL 448
 Db 381 PDSRRLRYRNKLGQFPYSIDLSNPQIPLIFG-----PIDAEDKMDTAMGYDYSNV 435
 QY 449 ADLISEPTOMPICVGLYAQMSGKSPFLKLEDEKKTAGQOTEPL-FQPSMLIVELTL 507
 Db 436 ADIVCEPESLPLTIGLYAKWGSKSALLAKLEAMHSRDMLDGVSLSVSPALFALE 495
 QY 508 ILGGLGLVFAFPV---DTMLAISTSPALTYIF---IYVFGRRGEGESNMWMA 560
 Db 496 LFGMEFLITMLLAISNTATALLISVYLLIFITFCSLIYVYVYGDGK---NWTSD 552
 QY 561 LSTRLAHIGYLELEFKLMFVNPPELPQOTKALPVFLTYDNRILSSVGGESTLAEMTA 620
 Db 553 IANFAVAFSRIRLVYNIITLHAP--MNEBDSAMPSVFLFADYHRLSSIGGEQALAKYA 611
 QY 621 TLSDACREGEPLATRLFRFRREDESGKKKKKKTKCCLPSFV-----IFLVYGCIIAGI 675
 Db 612 TLFEAAHTHGIVLPRFCGCKMPYPGIGHSLRHHGVPVILLIYAVFLIIMAOVGTIV 671
 QY 676 TLAIFRVDPKHLTVANILISVYGLAVLNCRTMVOYDLSLSQRRRLHVAASKLH 735
 Db 672 WILS--DRODNNFN--LFAIAPFLCGFVYAIYPLALIIYMTWVPPRRVRAAARNH 726
 QY 736 KLSSEGFMYLKCCEVELMARMATIDFTONOTRLVYIIDGLDACQDQYVQLMDTRVYL 795
 Db 727 KLRFEGLMOKLOTEVLDLADMIRSDFTRSHRLVYVVDGDNCEQEMVOTIDLETL 786
 QY 796 FS---KGFLAIFASDPHIIKAINONLSVLRDSINGHDIYRNIVHLPVFNLSRGLSN 852
 Db 787 FSARKHREFITIIAVDPHVIVSAIINMHMSALSCTELTGHDYKNTIISMPFYLHNSALKQ 846
 QY 853 ARKFLVTSATNGDITGSDTGTQEDTDRRYSONSLGEMTKLGSKALNRDITV----- 905
 Db 847 LQSKL-----RKRESMAEW-----KEKFKQDFFYGSHTSL 878
 QY 906 -----RRQORITRQMS-----FDITKLVTEDWFSDISPQTMRLNIVSVTGRLLR 955
 Db 879 READGRYSRKSSTVPNNKNSVYGRMMNGOILGEDYFSMNPBRAMRIYVALTLGRLMR 938
 QY 956 ANOITFMDWDLASMINTEOMPYRTSMLILYLETBGLPQOMLKTMYERISKNIPPTDQ 1015
 Db 939 AFETIDTSWMSLGHVWSTLEOWPFRKCMILLDRALEVNN--NQLLSVYIQLDHDIPADQD 996
 QY 1016 VEPLEIDGDIRNEFEVLSR---TPVLVARDVKTFLPCTVNLDPKLEIIT-----A 1064
 Db 997 ---LMDQDRNLLENFEGFLSKGIPSAERLTVGHVKKFVCTSLDYLKRLIRENSKGLV 1053
 QY 1065 DVRAARQINIGLAPPLPLHGGPPRPSGYSQPSVSSASFSNCPFFGGVYVSPQPHS 1124
 Db 1054 DIERA---QVGSAGMAIAP-----NARLLFSDDL----- 1079
 QY 1125 YSGLSGQPHFPPNRAAVPATGSSLLSSWTVDVYVEKROJEGDQDNNMPQYCTTIKA 1184
 Db 1080 -RMSIDTP-----LVKKLDVAVVNLIRKID-IPSNRLDSILDRFFYOL 1119


```

Db      154 OGHENVAHLINYGCKVRLPALHIAARNDTRPAVILLNDPNPDLVSKTGTPPLHIA 213.
      84 -----IHVEILLKSGASLEHRDMGWTALMACYKGRDVEILLSHGANPSYGLQYS 138
      214 AHEYNLNAQILLNGASVNTFPONGITPLHIAKRGVIMVRLLDGAGQLE-1K7KDE 272
      139 VPIITWAGRGHADIYHLLONGAVKNSDKYGTTPVWARKHLECVKHLANGADV 198
      273 LPLHCAARNGHVRISEILLDHGAPLOAKTKNGLSPIHMAAGDHLDICRLLLODAEID 332
      199 -----NSMTALYAVAGGTIOVKELK 225
      333 DITLDHPLHVAACGHHRAKVLDDKAGKNSALNGFPLHIAKKNHVRWELLK 392
      226 RNPVNLDDKDNATLMIASKEGHEIYODLADGTYVNIIPDRSGDYVLIGAVRGHEI 285
      393 TCASIDAVTESGLTFLHVASEFNGHLPIYKNLLORASVNSVAKVETPLHMAARGHEV 452
      286 VALLQKVDIDIRODNKATLYAVVEKGNATVPDIIQCNPDYEICTKDEPLIKATK 345
      453 AYTLLONKAKVAKAKDDQTPLCARIGHTMVKLENNANPMLATAGHTPLHIAAR 512
      346 MNIEVVELLDKGAKVASVDKDDTPLVHAIKGRSRLAELLBNPKDGLLYRPNKAG 405
      513 EGHVETVALLEKESQACMTKGGTTPLVHAKYKVAELLE--RDA-----HPNAG 566
      406 E--TPYNDISHQ-----KSL-----TQIFGAR----- 427
      567 KNGTLPPLHVAHHNMLDIYKLLPRGSGPHSPAMNGYTPHIAKONVEVARSLQYGG 626
      428 -----HSPETDGMIGYLYSALADITSEPLMOP-----PIC-- 462
      627 SANASVOGVYPLHIAEGHAENYALLSKQANONKNGSLPLHVAEGHVPADV 686
      463 -----VGLYACWGSK--SFLKCLUED--EMKTFAGQTEPLEFQSWL 501
      687 LKHGVMDATTRMGTPLHVAHSHGNIKYKFLLOHADVNAKTLEY--SPHQAQOQ 744
      502 --IVFLLLCGGLVPAFPD-----TNLAIASLPLALYIFPIYIFGGRGE 553
      745 GHTDVTLLKNG-----ASPNEVSDGTPLAIKRLGYISMDVLKVY----- 789
      554 SMNMAMALSTRARHIGYLELLFKMFVNPPELPEQTALPAVFLFDYKRLS-SVGE 612
      790 TDESFLVSDKNR--MSFPEVDELIVSEDEBELLS-----FKAERDRDDEE 840
      613 TSAEMIATLSDACEREFGLATRLR-----VPTEE--SOGKKKKKCTCLPSFV 662
      841 KELDFVPLQOYVE-----SPAIPRIKAMPETVYIRSEQOASKEVDEDSLIFS-- 892
      663 IFLFIVGCIAGITLLAIFRVDPKHLVYNAILISIASVGLAVLNCRTWMQVLDLILNS 722
      893 -----SPATEISDNI--SPVASPVHNGFLVSF-----AVDANGGS 925
      723 QKRRLHSAASKLHKLSKSGEMKYLKCEVELMARAKTIDSTQOQTRLYVITDILACEQ 782
      926 MGRSRHN-----GLRVVP-----PRCAAPTRITCRYL----- 954
      783 DKVLQMLDTPVLEFSKGFPIAFASDPHIIKAINONLSVLRDSNNGHDYMN--IVHL 841
      955 --KQKLTSTPPL-----AEEGLASIKIILAGPFGAOLFSPVIVEI 993
      842 PVFL--NSRGLSNARKFLVTSATNGDITCSDTTGTQEDTDRKVSNSLGEMTKLGSKTALN 900
      994 PHFASHGRG--DRELVIRSENGSV-----WKHEBSRYSYSYDQILN--GMBDELG 1041
      901 RRDYRRRQOQRTTRQMSFDULTLYT--EDMFSDISQPTMARLLNIYVYGRLLIRAN 957
      1042 SLEIEKKRVCRIIT--TDPLFYVIMSRICQD--YDITIGEG--GSLSKVLPLV 1091
      958 QITENMDRLASMINLTQOWPYRTSMLLYLEETGLDQOMTK-----TWYER 1005

```

```

Db      1092 QATPENAVTKRYKALQ-----AQPVDDELVTKLGNQATFSPITVEBR 1137
      1006 ISK-----NIPTTKDVEPLEIEDGIRNEVF-----LSRTPVIVA 1042
      1138 RRRHRRIGRLRPLRPWTONTNPRDSGSDTTSRLCSYGVGDQOAMEDIITTTLLVVA 1197
      1043 RDKTFELPCYVNIIDPKLREIIADVRAAREQINIGLAVPPLPHEGPPRRPSGYPASV 1102
      1198 NECANF--TTNV--SARFVLSDCPRTAEAVNPATILLYKEL-----TAVPYMAKF 1242
      1103 CSSASFNGPFPFGVSPQPHSSYSSGLSGQHFFYRBAAVPATGSSLLSMFVYDCEK 1162
      1243 VIFAKMDPREG-----MANNFGD--WHLFRSVYLEMKRSVESQVYPEDPR----- 1269
      1163 LRQ-----LDONMPOYCTTIKCAN-----INGRVL 1191
      1270 LEQHENVEVARSDIEVLEGMSLFAELSGNLVP-----YKKAQORSHFQSFRENRLA 1324
      1192 SOCNIDELKE-----MANNFGD--WHLFRSVYLEMKRSVESQVYPEDPR----- 1233
      1325 MPYVVRDSSREPGSLSFLRKAKMYEDTQHILCHLNTMPCKAGSGAEDRRTPPLAL 1384
      1234 ---FLNENSSAPYPHGSAR-----SHHELPLTELSTQPTTLNFSFEELNTGLD-- 1283
      1385 RYSLISESTGSLSTGEQAEKMAVISEHLGLMAELARE--LOFSVEDINRIKREMP 1440
      1284 ---EGAPRHSNLSWQSOTRRTPLSLINSODSIEISKLTDKVOA-----EYND 1329
      1441 NSLLEQSVALLNL--WVIREQONAMENITVALOSIDRGEIVNMLBESSQSRMLKDRKH 1499
      1330 AYREYIAQMSQLEBGTGSSSTISGRSSPHSTYYIGSSSGSISHSTLEDERGEGELKQED 1389
      1500 TDRDYSLSPQOMNG--YSSLQDELILSP-----ASLQALSSPLRADQ----- 1539
      1390 GRSFLMKRGDVIDYSSGVTNEASPLDPI--TEED--EKSQOS--GSKLIPGKS 1440
      1540 -----YNEVAVLDAIPLAATHEHTMLEMSDMQVSGALPGLVT 1579
      1441 SERPSLEFONDILKGGGLRYOKLPDSDESGTGRVQITPHCSKMITFKRLAKKORECASP 1500
      1580 AEDSSL-----ECSKAEDSDATGH-----EMKLEGLALSEPRRP 1613
      1501 QEHSAPETRFITAKLEYLDA--LDDKDDSDGCVGRSNSSPNHSLNEAADDSOLEKA 1557
      1614 ELGSLF--LYEDDTPVSDATNGTIDLE--QEEGGRSEKLPGRKODATGAGQ--DSE 1667
      1558 NLELEDEGHSGRGMPHSLSGLOPPIIARMSICSDK-----KSPSECSLIASSPESW 1612
      1668 NEVSIVSGHORGQARITHS-----PTVSQVTERSDRLQDMDADGSIVSYLDAAQGSW 1721

```

RESULT 5
AE2149
hypothetical protein al12748 [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp.
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
C:Accession: AE2149
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kunitz, T.; Sasamoto, S.; Watanabe, A.; Irigu
Nakazaki, N.; Shimizu, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AE2149
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-426 <CUR>
A:Cross-references: GB:BA000019; PIDN:BA074447.1; PID:q17131841; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: al12748

Query Match 6.3%; Score 555; DB 2; Length 426;

Best Local Similarity 37.2%; Pred. No. 3.3e-25;
Matches 149; Conservative 67; Mismatches 177; Indels 8; Gaps 6;

QY 9 VIVYEBENIPALKALLECKDYDERNEGQTPMLAEOGNEIYVELLKNGANCNL-E 67
Db 8 LKVAKSGDIKIGALLAGVGVICDRGTTALMNAIYGYEIVRSLLDGGANVNLAR 67
QY 68 DLONMWTALISASKEGHIHYEELLKSGASLEHRDMGWTALMACKGRDVEYLLSHG 127
Db 68 KRYGTRALMLAASANDVYDILLISRGAANATNEDSGTALMAALKGNVEARVALLAG 127
QY 128 ANPSVGLQYVPIIWAAGRGHADIYHLLONGAVNCSDKYGTTPLYWAARKHLECY 187
Db 128 ADVNITDKDDT-ALMLAYKRGQAAYVQIILLPSGADANCEDEGETLLMLAADSGHDYV 186
QY 188 KHLAMGADVDOEGASMTALYVAKGTYQSYVELLKRNPNVNLTDKGNALMASKE 247
Db 187 QVLLAGVGVNENOGGTALLAAGVANGALIKILLDRGADVNDHDDDESLALATYE 246
QY 248 GHIEIYODLLAGTYVPIIPRSQDPTVLIGAVRGHVEIYRALLQYKADIDIRGOD-NKTA 306
Db 247 GYDVYQVYVLLNGANVQIKKKLEDTPLLVAALOGHOIYETLL-KTG-ANVHGDNIGETP 304
QY 307 LYWAVERGNATVWDILQCPDTEICTKDETPILKATKRNIEVVELLDKAGKYSAYD 366
Db 305 LTLAASQGHATVRIILLDYGANANIPASDGTALIKATERNHGVIQILLAKGANVYQD 364
QY 367 KKGDTPHVAIRGSRRLAELLRNPKDGLLYRPKAGET 407
Db 365 SVGATALLMAASGGYNKVQIILLGADTNL--KNGGYT 402

RESULT 6

B35049

ankyrin 1, erythrocyte splice form 3 - human

N:Alternate names: ankyrin 2.1, erythrocyte; ankyrin-R

N:Contains: ankyrin 2.2, erythrocyte

C:Species: Homo sapiens (man)

C:Date: 17-Nov-1995 #sequence_revision 17-Nov-1995 #text_change 10-Jul-1998

C:Accession: B35049

A:Reference number: A35049; MUID:90175370; PMID:1689849

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1856 <LAM>

C:Genetics:

A:Gene: GDB:ANK1; ANK

A:Cross-references: GDB:118737; OMIM:182900

A:Map position: Bp11.2-Bp11.2

C:Superfamily: ankyrin; ankyrin repeat homology

C:Keywords: alternative splicing

F:2-1856/Product: ankyrin 1, erythrocyte form 3 #status predicted <MA7>

F:2-1513,1676-1856/Product: ankyrin 2.2, erythrocyte #status predicted <MA2>

F:44-76/Domain: ankyrin repeat homology <AN01>

F:77-109/Domain: ankyrin repeat homology <AN02>

F:110-142/Domain: ankyrin repeat homology <AN03>

F:143-171/Domain: ankyrin repeat homology <AN04>

F:172-204/Domain: ankyrin repeat homology <AN05>

F:205-237/Domain: ankyrin repeat homology <AN06>

F:238-270/Domain: ankyrin repeat homology <AN07>

F:271-303/Domain: ankyrin repeat homology <AN08>

F:304-336/Domain: ankyrin repeat homology <AN09>

F:337-369/Domain: ankyrin repeat homology <AN10>

F:370-402/Domain: ankyrin repeat homology <AN11>

F:403-435/Domain: ankyrin repeat homology <AN12>

F:436-468/Domain: ankyrin repeat homology <AN13>

F:469-501/Domain: ankyrin repeat homology <AN14>

F:502-534/Domain: ankyrin repeat homology <AN15>

F:535-567/Domain: ankyrin repeat homology <AN16>

F:568-600/Domain: ankyrin repeat homology <AN17>

F:601-633/Domain: ankyrin repeat homology <AN18>
F:634-666/Domain: ankyrin repeat homology <AN19>
F:667-699/Domain: ankyrin repeat homology <AN20>
F:700-732/Domain: ankyrin repeat homology <AN21>
F:733-765/Domain: ankyrin repeat homology <AN22>
F:766-798/Domain: ankyrin repeat homology <AN23>

Query Match 6.2%; Score 553; DB 2; Length 1856;
Best Local Similarity 20.5%; Pred. No. 5.3e-24;
Matches 388; Conservative 287; Mismatches 640; Indels 574; Gaps 73;

QY 21 LKALLECKDYDERNEGQTPMLAEOGNEIYVELLKNGANCNLEEDLONMWTALISASK 80
Db 94 VELVYNGANVNAQSQKFTPLXYMAOENHLEVVYFLELNGANQVATEDFTPLAVLAQ 153
QY 81 EGH----- 83
Db 154 OGHENVVAHLINYGKGVKRLPALHIAARNDDRTAAVLLONDPNDVLSKGTGTPHIA 213
QY 84 -----IHAYEELLKSGASLEHRDMGWTALMACYKGRDVEYLLSHGANPSVGLQYS 138
Db 214 AHYENLNAQOLLNNGSSVFTPPQGTTPHILASRGVNIYVRLLDGAGIE-TKIDE 272
QY 139 VYPIIWAAGRGHADIYHLLONGAVNCSDKYGTTPLYWAARKHLECYKHLAMGADV 198
Db 273 LPLHCAARNGHVRISEILLDHGAPIOAKTKNGLSPIHMAAGDHLDVRLLOYDAEID 332
QY 199 -----QEGA-----NSMTALYVAKGTYQSYVELLK 225
Db 333 DITLHPLHVAHAGHHRAKVLILLDKAKNSPALNGFTPLHACKNHRVWELLK 392
QY 226 RNPVNLDRKDGNTALMASKEGHIYODLLAGTYVPIIPRSQDPTVLIGAVRGHVEI 285
Db 393 TGASIDAVTESGLTPHVASFMGHLPYKNLLQRBASPVSVKYTEPLHMAARAGHVEY 452
QY 286 VRLAQKYADIDIRGDKNTALYWAVERGNATVWDILQCPDTEICTKDETPILKATK 345
Db 453 AKYLLONKAKVNAKKDDQTPHCAARIGHTMNVLLLENNANPILATTAGTTPHIAAR 512
QY 346 MNIEVVELLDKGVASVADKKGDPTPLHVAIRGSRRLAELLRNPKDGLLYRPKAG 405
Db 513 EGHVEVLLLEKESQACMTKRGFTPLHVAKYGRVAELLE--RDA--HPNAG 566
QY 406 E--TPYNDICSHQ-----KSLP-----TOLFGAR----- 427
Db 567 KNGTTPHVAVHNNDIYKLLPRGSGSPAMNGYTPPLHIAQONQYEVARSLOYG 626
QY 428 -----HLSPTEDDMIGDLYSSALADISEPTMOP-----PIC-- 462
Db 627 SANAESVQGVTPPLHIAOEGHAEVALLLSKQANGNLGKSGITPLHVAOEGHVPADV 686
QY 463 -----VGLYAOWGSGK--SFLLKLEED--EMKTFAGQGTPELPQFQSWL 501
Db 687 LKHGVMDATTRMGYTPPLHVASHTGNILKVLFLHQADVNAKRLG--SPLQAQO 744
QY 502 --IVFTLLLCGLGLVFAFPVD-----TNLAIASISFLAIYFFVIYFGRRRGE 553
Db 745 GHNDIYTLILKNG-----ASPNEVSSDGTTPPLAIKRLGISVTVLKVY----- 789
QY 554 SMMWMASTRLARHIGYLELLFKLMFVNPPELPEQTTKALPRLFTDYNFLS-SVGE 612
Db 790 TDETSVLVSKDHR--MSPEVTEDEILDVSEDEGEELIS-----FKAERDSRDVDE 840
QY 613 TSAEMIALTLDSCAEHFGFLATRLR-----VFTEP--SOGKKMKKTKCLSPV 662
Db 841 KELLDVVPKLDVVE-----SPALRIPCAMEPETYVNISEEOQSKTEDESLIPS-- 892
QY 663 IFLEIVGCIAGITTLAIFRVDPKHLTVNALISIASVGLAVLNCRTMVOVLDLSLNS 722
Db 893 -----SPATEFSQNI--SPVAASVPHGTFVSF-----MYDARGSS 925
QY 723 QRRRLHSAASKLHLKLSGFMVYLKCEVELLARMMAKTIDSFQONQTRLVYITDGDACEQ 782

```

Db 926 MGSRHN-----GLRVIP-----PRTCAAPRTICRLV----- 954
QY 783 DAVLOMLDITVRLVFSKGPFIAPFASPHIITIKAINONLSVYRDSNINCHDMRN-IYHL 841
Db 955 --KPOLSTPPPL-----AEEGLASRIITAGLPCTAOPLSVIYEI 993
QY 842 PVFL-NSRGLSNARKFLVTSATNGDITCSDTGTGODTRRVSNSLIGEMTLGSKTALN 900
Db 994 PHFASHGRS---DRELIVLRSENGSV-----WKEHRSRGESYLQIILN-GMDELIG 1041
QY 901 RRDYTRRQMOQTTRQKSFJTLKLVY---EDWESDISPQTRRLINIVSYTGLRLAN 957
Db 1042 SLEELKFKVCRHIT--TDFPLYFVIMSRLCOD-YDILCPeg-----GSLKSLVPLV 1091
QY 958 QITFNMRLASMINLLEQWPRYSMLILYLETBGLPQOMTLK-----TWER 1005
Db 1092 QATFPENAVTKKVALQ-----AQVPPDELVTKLLGNATSPITVPEPR 1137
QY 1006 ISK-----NIPTKDVEPLLEIDGDIRNEVF-----LSSRTPLVA 1042
Db 1138 RRRFHRPGLRILPSPMTDNFRDGEEDTSLRLCSVIGTDOAQMEDITGTLKLYA 1197
QY 1043 ROVKTFPLCTVLDKRLRITADVRAARQINIGLAPPLPLHGGPPRPSGVSQSPASV 1102
Db 1198 NECANF---TTNV--SARFWLSDCPRTAEAVNFATLLKEL-----TAVPYMAKF 1242
QY 1103 CSSASFNGFPFGGVVSPQPHSSYSGLSGOPHPFNRAAVPATGSSILLSMTVYVCEK 1162
Db 1243 VIFAKMNDPREG-----RLKCYCMTDKVKT 1269
QY 1163 LQOIGELDONMPPQYCTTKK---ANINGRVLSONIDELKEMANFGDWHLERS--M 1216
Db 1270 LEQHNFEVARSRIEVLGMSLFAELSG---NIVPYKKAQOGRSFPHQSPRENRL 1323
QY 1217 VLEMSVESQVVP-----EDPRFLNENSSAPVP-----HEBSARSRHTLPL 1258
Db 1324 AMPVAVRDSRREPGSLSFLKAMKYEDTQHLCLNLTTPCAGKSAEDRRRTPTPLA 1383
QY 1259 L--TELSQTPYTLNFSFE-ELNTLGLDEGAPRHSNLSMOSQTRTP--SLSLNSQDSSI 1314
Db 1384 LKYSILSESTPGSLGTDQAEKMAVISE---HLGSMALARELOPVSVDIN----- 1433
QY 1315 ELSKLDKVAQRYRAYREYIAQMSQLEGGSSTISGRSSPHSYTYI--GSSSSGGSIH 1372
Db 1434 -----RIWENPNSLLEQSVALLNL-----WVIRECONANMENILX 1468
QY 1373 STLEO-ERKKEGELKOEDEGRKSFILAK---RGVDYDSSGYSTNEASPLDPTTEDE--- 1425
Db 1469 TALOSIDNGEIVNMLEGSGROSRNLKPKDRHTDRYSLSPSQMANGYSSL-----QDELLS 1523
QY 1426 -----KSDQSGSKL-----LPKRSSEPSLFOQDLKLGGLRYOKLPSPDED 1468
Db 1524 PASLICALSPLRAQYVNEVAIILDAIPLAETHETMLEMSDMQVWAGL--PSLVAED 1582
QY 1469 ESGTRVOITPHCSKMITKRLKAKQRCASPOEHSAPITFTIKAKYILSDALLDKDS 1528
Db 1583 SS-----LECSR---AEDSDATGHEMKLEGALSEEPREPELSELVEDDVTSDAT 1631
QY 1529 S-----DSGVRNESPNSHLSHNEAADDQLEKANLILEDEGHSGKGMPSHLSGL 1580
Db 1632 NGLIDLLEQEGOREEKLPGSKRQDDATGAGQ--DSENEVSLVSGHQGQARITHS---- 1686
QY 1581 QDPITARMSICSEDK-----KSPSECSLIASSPEBSW 1612
Db 1687 --PTVSQVTERSODRLQDMDADGSIYSYLQDAAGSW 1721

```

RESULT 7

A35049
ankyrin 1, erythrocyte splice form 2 - human
N:Alternate names: ankyrin 2.1, erythrocyte; ankyrin-R
C:Species: Homo sapiens (man)

```

C:Date: 27-Jul-1990 #sequence_revision 01-Oct-1992 #text_change 04-Sep-1998
C:Accession: A35049
R:Lambert, S.; Yu, H.; Prechal, J.T.; Lawler, J.; Ruff, P.; Speicher, D.; Cheung, M.C.
Proc. Natl. Acad. Sci. U.S.A. 87, 1730-1734, 1990
A:Title: cDNA sequence for human erythrocyte ankyrin.
A:Reference number: A35049; MUID:90175370; PMID:1669849
A:Accession: A35049
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1880 <LAW>
A:Cross-references: GB:M28880
C:Genetics:
A:Gene: GDB:ANK1; ANK
A:Cross-references: GDB:118737; OMIM:182900
A:Map position: Bp11.2-Bp11.2
C:Superfamily: ankyrin; ankyrin repeat homology
C:Keywords: alternative splicing; cytoskeleton
F:2-1880/Product: ankyrin 1, erythrocyte form 2 #status predicted <MA1>
F:2-1513,1676-1880/Product: ankyrin 2.2, erythrocyte #status predicted <MA2>
F:44-76/Domain: ankyrin repeat homology <AN01>
F:77-109/Domain: ankyrin repeat homology <AN02>
F:110-142/Domain: ankyrin repeat homology <AN03>
F:143-171/Domain: ankyrin repeat homology <AN04>
F:172-204/Domain: ankyrin repeat homology <AN05>
F:205-237/Domain: ankyrin repeat homology <AN06>
F:238-270/Domain: ankyrin repeat homology <AN07>
F:271-303/Domain: ankyrin repeat homology <AN08>
F:304-336/Domain: ankyrin repeat homology <AN09>
F:337-369/Domain: ankyrin repeat homology <AN10>
F:370-402/Domain: ankyrin repeat homology <AN11>
F:403-435/Domain: ankyrin repeat homology <AN12>
F:436-468/Domain: ankyrin repeat homology <AN13>
F:469-501/Domain: ankyrin repeat homology <AN14>
F:502-534/Domain: ankyrin repeat homology <AN15>
F:535-567/Domain: ankyrin repeat homology <AN16>
F:568-600/Domain: ankyrin repeat homology <AN17>
F:601-633/Domain: ankyrin repeat homology <AN18>
F:634-666/Domain: ankyrin repeat homology <AN19>
F:667-699/Domain: ankyrin repeat homology <AN20>
F:700-732/Domain: ankyrin repeat homology <AN21>
F:733-765/Domain: ankyrin repeat homology <AN22>
F:766-798/Domain: ankyrin repeat homology <AN23>

Query Match 6.2%; Score 553; DB 2; Length 1880;
Best Local Similarity 20.5%; Pred. No. 5.4e-24;
Matches 388; Conservative 287; Mismatches 648; Indels 574; Gaps 73;

QY 21 LKALIECKDYDERNECGTPTLMLAABQGVETVELKLNKANGCNLEDDNMTALISAK 80
Db 94 VRELVTNGAVNMAOSOKGFTPLVMAOENHLEVKFLENGANQNVATEDGETPLAVALO 153
QY 81 EGH----- 83
Db 154 QGHENVAHLINLTGKRVRLPALHTAARDRTAAVLQNDPNDVLSKGTPTPLHTA 213
QY 84 -----IHIVELLKSASLEHRDMGWTALMAVCYGRDYVELLSHGANSPTGLQYS 138
Db 214 AHENENLVADQLINRRESSVFTPONGITPLHIASRGNVYWRLLDRAQGLE-TRKDE 272
QY 139 VYPTIYAAGRGADYHLLQNGAKVCSDKYTTPPLVMAARKGHIECYKHLAMGADV 198
Db 273 LTPHCAARGHVRHISEILLDRGAPLOARTKNGLSPIHMAAGDHLDVRYLLQYDAEID 332
QY 199 -----QEGA-----NSMTALIVAKGYTQSVKEILK 225
Db 333 DITLDHLPLTHVAHGHGHRVAVKVLDDKAKRPSRLNFTPLHACKNHRVYVELLIK 392
QY 226 RNPVNLTJDKDGNTALMAISKEGHIEYODLDAGTYVNI PDRSGDYVLIGAVRGHVEI 285
Db 393 TGSASIDAVESGILPLHVASFMCHLPIYKNLLQRGASPVSVNKKETPLHMAARAGHTEV 452
QY 286 VRALLQRYADIDRGDNKTALYMAVEKSNATVWRDILQCPDTEICTKDGTEPLIKATK 345

```

```

Db 453 AKYLONKAKVAKAKDDQPLHCAARIGHTNMVKLLENNANPNLATTAGHTPLHIAAR 512
QY 346 MNIEVVELLDKGAQVAVDKKGTPLHVALRGSRLAEILLNPKDGLLYPNKAG 405
Db 513 ECHVEFVALLEKESQACMTKGTPLHVAKYKRVALLLE--RDA--BPNAAG 566
QY 406 E--PPYNDICSHQ--KSL--TOFGAR-- 427
Db 567 KNGPLPLHVAHHNNLDIYKLLPRGSGSPHAMNGYTPPLHAAQONOEVAKSLLOYG 626
QY 428 -----HLSPTEDGMDLGYLSSALADISEPTMOP-----PIC-- 462
Db 627 SANAESVQGTPLHAAQGHAEVALLSLKQANGNLKSGLTPLHAAQGHAVADY 686
QY 463 -----VGLYAQMSGR--SFLAKLED--EMKTFAGQGTPLPOFSL 501
Db 687 LKHGVMDATRMGTPLHVAHNGNKLKVLKLOHADVNAKTKLG--SPLHQAQO 744
QY 502 --IVFTLLCGGLVFAFPVD-----TNLAISLSFLALITYFVIYFGGRGE 553
Db 745 GHRTDITLLKNG-----ASPNEVSDGTPPLAIAKRLGISTVTLKAV----- 789
QY 554 SNMAMALSTRLARHIGYLELLFKLMFVNPPELPGQTKALPYRELTDYNRLS--SVGE 612
Db 790 TDETSFVAVSDKHR--MSPEYDEILDVSEDEGELLIS-----FKAERDSRDVDE 840
QY 613 TSLAEIATLSDACEHEFGFLATRLR-----VFTEE--SQCKKKKTKCCLPSV 662
Db 841 KELLDVFPKIDOVE-----SPAIPRICAMPEYVISEEQEQAKEYDSDSLPS-- 892
QY 663 IFLFIYCIAGITLAIIRVDPKHLTVANAIIISIVGLAFVLCRTMQLDLSLNS 722
Db 893 -----SPATTSQNT--SPVAPYHTGTVSF-----MVDARGSS 925
QY 723 QRRRLHSAASKLHLKESGFMKYLKEVELAMAMAKTIDSTQONQRLVLIIDGLDACEQ 782
Db 926 MRGSRHN-----GLRVYIP-----PRCAAPRITCRV----- 954
QY 783 DKYLOMLDYRVLFSGKPLAIFASDPHIIKAINONLSVLRDSNINHDMYRN--IYHL 841
Db 955 --KPKLTSTPPPL-----AEEGLASRIIALGPTGAQFLSPVIVEI 993
QY 842 PVFL--NSRGLSNARKFLVTSATNGDITCSDTTGTQEDTRRVSQNSLGEKTKLSKTAALN 900
Db 994 PHRASGRG--DRELVLRSKESV-----KHEHSRGESESLDQILN--GMEELG 1041
QY 901 RRDYRRRQMRITTKQMSFDTLKLVT--EDWESDISQTMRLNLINVSYTGRLLRAN 957
Db 1042 STEELKKRVCRIIT--TDFPLFYVIMSRLCD--YDIIGPEG-----GSLKSKLVPLV 1091
QY 958 QITFNMRLASWINLEQWYRTSWLILYETEGLPDQWTK-----TMYER 1005
Db 1092 QATFPENAVTKRKLALQ-----AQPVPDELVTYLLCONATFSPITVEEP 1137
QY 1006 ISK-----NIPTKVDEPLELIDGIRNFEV-----LSSRPFLVA 1042
Db 1138 RRRFHRIGIRILPSPWMDNPRDSGDTSLRLCSVIGTDQAWEDITGTTLVYA 1197
QY 1043 RDKYTLPLCTVNDPKIRELIAVRAAREOINIGLAPPLPLHCEGPPRPSPSY 1102
Db 1198 NECANF--TTNY--SARFMLSDCPRTAEAVNFATLLYKEL-----TAVPYMAKF 1242
QY 1103 CSSASFNGPRPGVAVSPQHSYVSGLSGQPHFYNAAPVATGSSILLSSMVDVACEK 1162
Db 1243 VITAKAMDPRG-----RLKRCMTDDKVDKI 1269
QY 1163 LROLEGIDQMMPOYCTTIK--ANINGRLVSOENIDELKEMANFEGDWHLFRS--M 1216
Db 1270 LEQJENFVEARSHDIEVLGMSLFAELSG-----NLVYVKAQAQOORSHFQSPRENRL 1323
QY 1217 VLEKRSVESOVVP-----EDPRLNENSAPVP-----HGEARSSSHTELP 1258
Db 1324 AMPYKVDSSREPQGSLSPLRKAKMYEDTOHILCHLNTMPKCAKSGAGDRRRTPPLA 1383

```

```

QY 1259 L--TELSQPTPLNFSP--ELANTLGLDEGAPRHSNLSMOSQTRRP--SLSSINSODSSI 1314
Db 1384 LRYLSLSTSTGSLSGTEQAEKMAVISE-----HGLSNMELARELOFVEDIN----- 1433
QY 1315 EISLTKVQAEYDAYREYIAOMSOLEGTGSSFTISGRSSPHSTYI--GSSSSGSIH 1372
Db 1434 -----RIVENPNLSLEQSVALLN-----WVIREQANAMNENY 1468
QY 1373 STLEO--ERKRGELKOEGRKSFMLK--RGDVIDYSSGVSTNEASPLDPIREDE-- 1425
Db 1469 TALQSIDGELVNNMIEGSGROSRLMKPDRRTDSDYSLSPQMGYSYL-----QDELS 1523
QY 1426 -----KSDQSGSL-----LPGRKSSERPLFTDLKLGKGRYOKLPDESD 1468
Db 1524 PASLQALSSPLRADQYNEVAIIDAIPLATIETHDMLKMSQDMQVNSAGLT--PSLYTAED 1582
QY 1469 ESGTGRVOITPHCSMIRTKRLKAKKORCASPOEHSAPPIRTIKAEYSDALLDKDS 1528
Db 1583 SS-----LECSK--AEDSDATGHEWKLLEGALSEPRGPELSLEVEDDVTDSAT 1631
QY 1529 S-----DSGVSNESSPHNSLHNEAADSQLEKANLILEDEBGSKGMHSISGL 1580
Db 1632 NGLIDLBOEGQREKRLPGSKRODDATGAGQ--DSENEVLSVGHORGARLTHS----- 1686
QY 1581 QDPIARMSICSEK-----KSPSECSLIASSPEESW 1612
Db 1687 --PTVSQYTERSQDRLOQMDADGISTVYSLQMAQGSW 1721

```

```

RESULT 8
T42716
Ankyrin 3, splice form 4 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000
C:Accession: T42716
R:Peterson, L.L.; John, K.M.; Lu, F.M.; Elcher, E.M.; Higgins, A.; Yialamas, M.; Turtzo
J. Cell Biol. 130, 313-330, 1995
A:Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin ge
the repeat domain.
A:Reference number: 222237; MUID:95340633; PMID:7615634
A:Accession: T42716
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1961 <PEP>
A:Cross-references: EMBL:L40632; NID:6710548; PID:67110552; PIDN:AAB01607.1
A:Experimental source: strain C57BL/6J; kidney
C:Genetics:
A:Gene: Ank3
A:Map position: 10
C:Superfamily: ankyrin repeat homology
C:Keywords: alternative splicing

```

```

Query Match 6.2%; Score 552; DB 2; Length 1961;
Best Local Similarity 20.2%; Pred. No. 6.7e-24;
Matches 427; Conservative 289; Mismatches 674; Indels 724; Gaps 85;

QY 14 EENPILAKALLEKKDVERNEGOTPLMLAEOGNEIYKELKNGANCNLEDDIWT 73
Db 66 KEGHVEVVELDQREANDAATKGNATLHLSLAGQEVYKVLVTNCAVNAASQNGFT 125
QY 74 ALISAKEGHIIIVELELGSASLEHRDGGWTALMAACQYKGRDVELEL----- 124
Db 126 PLYMAQENHLEVVRELLDNGASQSLATEGDTPLVALVLOQCHDOVLSLELNDTKGVR 185
QY 125 -----SIGANPSYTG- 135
Db 186 LPALHIAARKDDTKAALLQNDTNADVESKGTPLHIAHYGININATLLNRAAIVD 245
QY 136 --OYSVYPTIWAAGGHADIYHLLONGAKVNSDKYGTTPPLVWAARKG----- 182
Db 246 FTARNDITPLHVAASKRGANMYKLLLDKGAIKADATRQGLTPLHGAASGHQVYEMLLD 305

```

QY 183 -----HLECYHLLAMGADVDEGANSMTALIVAKGTYOS 219
 Db 306 RSAPILSKTKNGSLPLHMAOGDHLNCGVOLLQHNVPDDVTNDLTLALHVAACHGCHKV 365
 QY 220 VAEILKRNPNVLTJDKCNTALMIASKE----- 247
 Db 366 AAVLLDKRSPAKALNGFTPLHIAKCKRIRIYVMEILLKHGASIQAVTESGLTPIHVAAF 425
 QY 248 -GHIEYODLLDAGTYVNIPIRRSGDVLIGAVRGHVEIVRALLQKYADIDIRGDNKTA 306
 Db 426 KGHVNIIVSGLMHGHSPTNTYRGETALHMAARSOAEVRYLVODQAVEKAKADDTPI 485
 QY 307 IYWAIVEGNATVNDILOCPNDEICTDGETPLIKATKRNIEVEVLELLDKGAVSYVD 366
 Db 486 LHISARLGKADIVQOLLOOGASPNMATTSGYTPLHAAAREGHEVAAFLDDHGSLSLTT 545
 QY 367 KCGDPTLHVAIRGRRLAEILLR-----NPKGRLL----- 398
 Db 546 KKGFTPLHVAARYKLEVASILLQKSASPDAGKSGTLPPLHVAHYDQKVALLLDDGA 605
 QY 399 --YRPNKAGEPPYNIIDC-SHOKSILTOI--FGAR-----HLSPTETGDMGY 441
 Db 606 SPHAAKNGYPLHIAAKKNQMDITSLIEYGADANAVTROGIASVHLAABGHDVAVSL 665
 QY 442 DLYSALADILSEPTEMOP-----PICVGLYAWQ 469
 Db 666 LLSRANAVNLNKSGLTPLHAAQEDRYNAEVLVNOGAHVAQCKMGYTPLHVGCH--Y 723
 QY 470 GSGK--SFLK--KLEDEMKTFAGQTEPLFQFSWLVFLTL-----LLCGGLGV 516
 Db 724 GNIKTVNLLOHSAKAVNAKTKNGYALHQAQOQHTHITLVLLONNASPNELTVNG-- 779
 QY 517 FAFVDTMLAIALISLFLALYIFPIVI----- 544
 Db 780 -----NTALAIARLGIVSYVDLKVYTEEIMTTTTTEKHMANPETMNEVLDMSDEV 834
 QY 545 -----YGGREGESWMAALSTRILAHIGYELLFKL-----MFIN 582
 Db 835 RKASAPKLSOGEYISDEEGDKTWFKIPRYO-----EVLVYSDATITGDTDYXIG 886
 QY 583 PPELPEQTTKALPVFLFTDYNRLLSSVGETSLAEMITLSDACEREGEFLATPLFRFR 642
 Db 887 PODKELGDDSLPAE-----GYVG--FSLGARSASLR-----SESSDSYTLNR 928
 QY 643 TEESQKRRKKMKYCCLPSFVIFLIVGCIITLALFRVDP--KHLT-----VNAI 693
 Db 929 SSYAR-----DSMITEELLVPSKEQHLFTFTREFSDSLRHSMADTLDNNVLY 977
 QY 694 LISTASVYGLAFVLCRTWQVLDLSLNSQ-----RK-----RLHSASKLHL- 737
 Db 978 SSPVHSGFLVSFVDARG-----GSMRGRHGHKRIIIPRCKTAPTRITCRLVKRRKLA 1032
 QY 738 -----KSEG-----FMKVLKCEYELMARAKTIDSTON 766
 Db 1033 NPPVAVBEGELASRLVEMGRGAQPLGIVYVEIRHFGSMRKEKELLIVLSENGETWKEH 1092
 QY 767 Q-----TRLVVIIDGLACEQDKVLOMLDVRV--LFSKGFIAFASDPHIIKAIION 819
 Db 1093 QPDSKNEDLAEILLNGMD--EELDSPEELGTFRICRIITK--DEPYFA-----VYSRIKOE 1144
 QY 820 LNSVLRDSNININGHYMINIYVLPFLMSRGLSNARKVLY--TSATNGDITSDTGTG-- 875
 Db 1145 SNOIGPBGCI-----LSSITVPLVYASPEDEGALKRIRIRVGLQAOPI 1184
 QY 876 --EDTRRSONSLGEMTKLGSKTALNRDPT--RRROMORTITROASFDLTKL--VT 928
 Db 1185 VPEETVAKI-----LGNKATFSPITYVERRRKKFKPIT--MTIYPPPSGGEGVS 1232
 QY 929 EDWFSDISPQTMRLNIVSVTG-----RLRANQITFNMDRLASWNLTEOWEYRT 980
 Db 1233 NGYKGDATPNL--RLT--CSITGTSPOAWMEDINGTPLPLFKDCVAFSTTVNSARF-- 1284
 QY 981 SWL-----ILVLEETBEGLPQOMILKTM--YERISKINITTTKDVPELLEIDSDIRNF----- 1029

Db 1285 -WLADGQVLL--ETVLSAQILRELICVYAKKVFVAKTNDP--VESLSCFCMTDD 1337
 QY 1030 -----EVLFSRTPVLVAVDVKTLFCTVINDLPKLR--EIIADYRAARE-- 1071
 Db 1338 RVDKTLBOQNFEEVARSKDIEVLEKPI--YVDCGNLAPLKGQOLVFNFSKERN 1395
 QY 1072 ---QINIGLAPPLP--LHGGPPRPSGYQSPASCSASFNPFGGVYSPQPHSS 1124
 Db 1396 LPSIKIRDSQECGRSLFLKE--PKTKGLPOTA--VCN----- 1432
 QY 1125 YVSGLSGPQHPFYRAAVPATGSSILL-----SMKTVDVYCEK--LRQIEGLDOMN 1173
 Db 1433 --LNTILPAKKKEKADRRSPASLAKRYSTLTPSPMSPOPCERTDIRALVAD--HL 1489
 QY 1174 MPQCTTIKKNANINGRYLSQCNIDELKEMANF--GDM-----HLFRSVLEBM 1220
 Db 1490 GLSWTELARLNFVDEINQIRVENPNLSISQSFMLKKWVTRDKNAATDALTSVLTKI 1549
 QY 1221 RVSQGVVPEDP-----RFLNENS--SAPVPHGSARSSHTLPLTELSSQTP 1267
 Db 1550 NRIDIVTLBGLPIFDYGNISGTRSFADENNVPDPV-----DGH--PSFOVELETP 1598
 QY 1268 YTLNFSPEELNTLGLDEGAPRHSNLSMOSQTRTPSLSS-----LN 1308
 Db 1599 MGLYWT--PNNPQOQD--HFSDISIESPFTPRLSDGLVPSCGNIEHPGPPVYT 1653
 QY 1309 SQDSSTELSKLTQV--QAETRDAYREYIAOMSQLBEGTSGSTISG--RSP-- 1356
 Db 1654 AEDTSLSDSKMDSVYTPDAPDLVDDESQKLQCECAQCAWVAGIIPNDGRQAEPLR 1713
 QY 1357 HSTYVIGOSSSGSISHTLEOERKEGELKQEDGRKSFIMKRGVDIDYSSGVSTNEASP 1416
 Db 1714 PQRKXGMSE-----QQEKXSG-----P 1733
 QY 1417 LDPITEDEKSDSGSKLLPGKKSSEPSLFTQDLKLGGLRYOKLPSPDEDESGTRVQ 1476
 Db 1734 DEEVTEDEKVK-----SLFE-DIOLE--EVEAEEMTEDEQOAMILNRVQ 1772
 QY 1477 ITPHCKMIRTKRLKAKORECASPOHSAEPIRTFKAEYLSDALDLKKKSDSGSVSN 1536
 Db 1773 -----RAELMSSLAQWNETPSSLSLEPAQAR-----LVLGDLDDSDOARDS 1820
 QY 1537 -----ESSPNHSLHNEAADDSOLEKANLIELEDEHSGKGM-----PHSL 1577
 Db 1821 ITSYLGEFGKILANNH-----AEVPEAKAKPYFPSSQNDIKGOSIKEMLKRTKHC 1875
 QY 1578 SGLQDPIIAMSICSEDKKSPSECS--LIASSPESWPAQORAYNLNRTPTVTLNNNTA 1635
 Db 1876 GRTEEPV--SPLTAYOKSLSETSKLVIEDAPKPCVPVGMKKM-----TFTT 1919
 QY 1636 PTRANONPDETEG 1649
 Db 1920 ADGKARLNLQEEEG 1933

RESULT 9
 T42715
 ankryn 3, splice form 3 - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000.
 C:Accession: T42715
 R:Peterson, L.L.; Johnson, K.M.; Lu, F.M.; Elcher, E.M.; Higgins, A.; Yiallams, M.; Turtzo
 J. Cell Biol. 130, 313-330, 1995
 A:Title: Ank3 (epithelial ankryn), a widely distributed new member of the ankryn ge
 the repeat domain.
 A:Reference number: Z22237; MUID:95340633; PMID:7615634
 A:Accession: T42715
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-1940 <P>
 A:Cross-references: EMBL:L40632; NID:g710548; PID:g710549; PIDN:AAB01604.1
 A:Experimental source: strain C57BL/6J; kidney

QY 1599 SECS--LIASSESWPACOKAYNLNRPSTVTLNNTPTNRANONPEDEIG 1649
 Db 1872 EETSKVIEDAPKPCVPVGMKK-----TRTADCKARLNQEEEG 1912

RESULT 10

T42713
 ankyrin 3, splice form 1 - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000
 C:Accession: T42713
 R:Peterson, L.V.; John, K.M.; Lu, F.M.; Elcher, E.M.; Higgins, A.; Yialamas, M.; Turtzo, J.
 J. Cell Biol. 130, 313-330, 1995
 A:title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin gene
 A:repeat domain.
 A:Reference number: 222237; MUID:95340633; PMID:7615634
 A:Accession: T42713
 A:status: preliminary; translated from GB/EMBL/DBJ
 A:molecule type: mRNA
 A:Residues: 1-1943 <PEP>
 A:Cross-references: EMBL:L40632; NID:g710548; PID:g710550; PIDN:AAB01606.1
 A:Experimental source: strain C57BL/6J; kidney
 A:Genetics:
 A:Gene: Ank3
 A:Map position: 10
 A:Introns: 855/1
 C:Function:
 A:Description: supposed to play an important role in the polarized distribution of many
 C:superfamily: ankyrin; ankyrin repeat homology
 C:Keywords: alternative splicing

Query Match 6.2% Score 551; DB 2; Length 1943;
 Best local similarity 20.2%; Pred. No. 7.5e-24;
 Matches 426; Conservative 289; Mismatches 666; Indels 724; Gaps 85;

QY 14 EENIPALKALIECKVDNERNEGOTPLMAEGGNEIYKELLKNGANCLELDLMMT 73
 Db 66 KEGHEVVESELLQREANVDAATKRGNTALHIASLAGQAEVVKVLTNGCANVNAQSNGFT 125
 QY 74 ALISAKGHIHIVELLKSGASLEHRMGNTALMAACYGRDVEYELL----- 124
 Db 126 PLYMAOENHLEVEFLLDNGASOSLATEDEFTPLAVALQGHQVYSLLENDTKKVR 185
 QY 125 -----SHGANPSVTGL----- 135
 Db 186 LPALHIAARKDPTKAALLLQNDINADVESKSGFTPLHIAHYNINATILLRAAAYD 245
 QY 136 ---QSYVPIIWAAGRGHADIYHLLONGAVNCSDKYGTTPLYWAARKG----- 182
 Db 246 FTARNDITPLHVASKRGANVVKLLDRGAKIDAKTRDGLTPLHCGASGHEQVEMLLD 305
 QY 183 -----HLECYKHLANGADVDOGASMSRALLIYAVNGGYTOS 219
 Db 306 RSAPILSKTKNGSLPLHATQGDHLCVQLLDHNNVPDVTNDYLTALHVAACHGKRV 365
 QY 220 VKEILKRNPNVNLTDKDGNTALMTASKE----- 247
 Db 366 AKVLLDKKASPNAKALNGFTPLHACKKNRIKRWELLKHGASTOAVTESGLTPIHYAIF 425
 QY 248 -GHEIYVDDLDAGTYNIPRSGDYVLIGAVRGHEIYRALLQYKADIDIRGODNKA 306
 Db 426 MGHVNIYQALMHHGASPNVTITVREGTALHMAARSGOAEVYRIVQDAQVAKAKDDQTP 485
 QY 307 LYMAVEKGNATVBDIIOCNPDTEICRKDGTEPLIKATKMNIEVEVELLDKGAAYAVD 366
 Db 486 LHISARLCKADIVQLLQOGASPNAAATSGTTPHLHARSGHEVAAFTLLDHGASISITT 545
 QY 367 KKGDTPHVAIRGSRRLAELLIR-----NPKDGRLL----- 398
 Db 546 KKGFTPLHVAIKYGLKLEVASLLDQKSPDAAGSGTLPLHVAHAHYNOAVALLLLDQGA 605

QY 399 -YRPNKAGENPVNIDC-SHOKSILNOL--FGAR-----HLSPETDGMAGY 441
 Db 606 SPHAAKNGTTPPLHIAAKKNOMDIATSLLEYGADANAATROGLASVHLAAQESHVAVSL 665
 QY 442 DLYSSALADILSEPTMOP-----PICVGLYAQW 469
 Db 666 LLSRNANVNLNKSGLTPPLHIAQEDRVNVAEVLVNGAHVDAQTKMGYTPPLAVGCH--Y 723
 QY 470 GSGK--SFLLK--KLEDEMKTRAGQOTEPPLFQPSWLYVTLT-----LLCGIGLV 516
 Db 724 GNIRIVNPLQHASKVAKTRNGYTALHQAQOGHTHIIIVLLQNNASPNELVNG----- 779
 QY 517 FAFVPDNMLIAISLALITYEIFYI----- 544
 Db 780 -----NHALIARLGLIISYVDILKYVTEIMTTTTEKKNVPEYMEVLDMSDEX 834
 QY 545 -----YFGRRGEESNMMAALSTRHAGVLELLKIMFVNPPELPEQOT 591
 Db 835 RKASAPKLSDEGEYISDGEED-----AITGDTKYLQ-----PQDLKELGD 877
 QY 592 KALPVRFLFDYNRSLSSVGETSLAEIATLSDACEREFGLATRLFRVFRTEESGKKK 651
 Db 878 DSLPAAE-----GVVG--FSLGARSASLR-----SFSSDRSYTLNRSYSAR-- 915
 QY 652 WKTCCLPSFVIFLEFYICIIAGITLLAIFRVPD-KHLT-----VNAILISASYVG 702
 Db 916 -----DSMIEELVPSKEQHILTFREDOSLNRHYSWADTLNVLVSPVHSGFL 968
 QY 703 LAFVLCRTWMOYVLDLSLNSQ-----RK-----RLHSAASKLHKL-----KSE 740
 Db 969 VSEKVDARG-----GSMRGRHGMRIIPRCKTAPTRITCLYVRHKLANPVPVVEGE 1023
 QY 741 G-----FMKYLKCEVELAMAMATIDSTFQNO-----TLU 770
 Db 1024 GLASRLVEKMPAGAQLFPIYVEIPHFGSRGRERELIYLRSENGETWEKHQSKNEDL 1083
 QY 771 VVIIDGLACEQDQVLOMLDTVVR--LFSKGPRIATFASDPHIIKAIQONLNSVYRDSN 828
 Db 1084 AELLNGMD--EELDSPELQTKRICHLITK-DEPQYFA-----VYSRIQESQIQPEGG 1135
 QY 829 INGHDIYKRNIVHLVPLNLSRGLSNARKFLV-TSATNGDITCSDTTGTQ-----EDTDRV 882
 Db 1136 I-----LSSTVPIVQASFEGALTKRIRVGLQAPVEETVKKI 1175
 QY 883 SONSIGEMTKLQSKTALNRDNY--RRROQRIITQOMSEFDLTKLL-----VTEDMPSDISP 937
 Db 1176 -----LGKKAIFSPPIYVEPRKRFHKFTI--MTIPVPPSGEGSVSGYGDATP 1223
 QY 938 QTRRLINIVSYTG-----RLRANOITFNMORLASIMLITEQWPIRTSML-----I 984
 Db 1224 NL--RLT--CSITGTSAPQWEDITGTPPLTFIKDCVSTTNVSARF-----WLADCHQV 1274
 QY 985 LYLEETEGPLDQMTLTM-YERISKNIPTTKDYVEPLLELDGDIRNF----- 1029
 Db 1275 L--ETVGLASQLYRELICVPYMAKFFVPAKTNDP--VESSILRCFCMTDDRVKTLDEQ 1328
 QY 1030 -----EYFLSRFVLAARDVKTPLPCTVNLDEKLR--EILADVAAARE-----QINIGG 1077
 Db 1329 ENEEYARSKDIEVLESGRI--YVDCYGLADPLTKGQGLVNFYSFKNRLPSSIKIRD 1386
 QY 1078 LAYPLP-----LHEGPPRPSPGYSOPASVSSASFNGPPGVSPOPHSSYSGLSGPO 1133
 Db 1387 TSQEPGRLSFLKE--PKTTKGLPQTA--VCN-----LWITPLA 1421
 QY 1134 HPPYNAANVPATGSSILL-----SMYVDVYCEK--LRQIGLDONMMPQYCTTTP 1182
 Db 1422 HKKAERADRQSPASIALKRYRYLTPESMSPQSCERTIDIMATVAD--HIGLSWTELAR 1480
 QY 1183 KANINGRVLSOONIDELKKEAMNF--GDM-----HLFRSVLEMRSVSSQYVP 1229
 Db 1481 ELNFSYDEINQIRVEPNLSISQFMALKKWTTRGKNAATTDALTSVLTAKNRIDIVTL 1540
 QY 1230 EDP-----RFLNENS--SAPVPHGESARRSHTLPLTELSSQTYTLNFSPEE 1276

Db 1541 EGIIPDYGNISGTRSRADENNHYHDV-----DGH-----PSFYQLELTPMKLYMT--P 1587
 QY 1277 LNTLGLDEGAPRHSNLSWOSQTRTPSLSS-----LNSODSIEIS 1317
 Db 1588 PNFPOQDD--HFSDISSIESPPRTSRLSDGLVPSQGNIEHPTGPPVTAEDTSLSDS 1644
 QY 1318 KLTDKY-----QAEYRQAYREYIAQMSOLEGCTSSITSG-----KSSP--HSTYYIGOS 1365
 Db 1645 KMDSDVTVTPDAPPLVDDESQKLDLCQSECAQCMASVPGIPNDGRAPLELRPQTRKVGMS 1704
 QY 1366 SSGSISHTLEDRGKEGELKQEDGRKSFIMKRGVDYSSGSVTNEASPLDPTIDEDE 1425
 Db 1705 SE-----QDEKGSNG-----PDEEYVEDKV 1724
 QY 1426 KSDQSGSKLLPGKSSSERPSLFQTDLKLKGGLRYQKLPSDEDESGTRGVQITPHCSKMI 1485
 Db 1725 K-----SLFE-DIQLE--EVEAEEMEDQOQAMLNRYQ-----RAEL 1758
 QY 1486 RTRRLAKAKORECASPOHSAEPTRTFKAKEYISDLMLDKKSSDSGSVRSN----- 1536
 Db 1759 AMSSLAGWONEPSPGSLSPQAQR-----LTGGLDLDDSDQARDSTSYLTGEF 1811
 QY 1537 ---ESSPNHSLHNEADDSQLEKANLIELEDEHSGKRGV-----PHSLSGIADPITA 1586
 Db 1812 GKIEANGNHT-----AEVPEAKAKAPYEPESQNDIGKQSTIKEMLKPKTHCGRTPEPV-- 1864
 QY 1587 RMCISCDKKSPESECS--LIASSPESWPACQKAYNLNPTSTVTLNNNTAPINRANONF 1644
 Db 1865 --SPLTAYQKSLSETSKLVIEDAPKPCVPVGMKRM-----TRTTADGKARLNL 1910
 QY 1645 DEIEG 1649
 Db 1911 QDEEG 1915

RESULT 11

37431

ankyrin 2, neuronal long splice form - human

N:Alternate names: ankyrin B, 440K splice form; ankyrin-B, brain ankyrin; non-eythroid

C:Contains: ankyrin 2, short form

C:Species: Homo sapiens (man)

C:Date: 06-Jan-1995 #sequence revision 06-Jan-1995 #text change 13-Aug-1999

C:Accession: S37431; A39643; B39643; A40334; A49462; S14533; S14569

R:Chan, W.

Submitted to the EMBL Data Library, September 1993

Reference number: S37431

A:Accession: S37431

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-3924 <CHAS>

A:Cross-references: EMBL:Z26634; NID:9406287; PIDN:CAA81387.1; PID:9406288

R:Otto, E.; Kunimoto, M.; McLaughlin, T.; Bennett, V.

J. Cell Biol. 114, 241-253, 1991

A:Title: Isolation and characterization of cDNAs encoding human brain ankyrins reveal a

A:Reference number: A39643; MUID:91302466; PMID:1830053

A:Accession: A39643

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-2077 <OT1>

A:Cross-references: GB:X56957

A:Accession: B39643

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1443,3585-3924 <OT>

A:Cross-references: EMBL:X56958

R:Tsai, W.T.; Menninger, J.C.; Yang-Feng, T.L.; Francke, U.; Sahr, K.E.; Lux, S.E.; Ward, Genomics 10, 858-866, 1991

A:Title: Isolation and chromosomal localization of a novel nonerythroid ankyrin gene.

A:Reference number: A40334; MUID:92009921; PMID:1833308

A:Accession: A40334

A:Molecule type: DNA

A:Residues: 463-474, 'PE', 477-495 <TSE>

A:Cross-references: GB:M37123; NID:9178647; PIDN:AAA62828.1; PID:9178648
 R:Chan, W.; Kordeli, E.; Bennett, V.
 J. Cell Biol. 123, 1463-1473, 1993
 A:Title: 440-kD ankyrinB: structure of the major developmentally regulated domain and
 A:Reference number: A49462; MUID:94075409; PMID:8235844
 A:Accession: A49462
 A:Status: preliminary; nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-3924 <RES>
 A:Cross-references: EMBL:Z26634; NID:9406287; PIDN:CAA81387.1; PID:9406288
 C:Genetics:
 A:Gene: GDB:ANK2
 A:Cross-references: GDB:127607; OMIM:106410
 A:Map position: 4q25-4q27
 C:Superfamily: ankyrin, ankyrin repeat homology
 C:Keywords: alternative splicing
 F:2-3924/Product: ankyrin 2, long form #status predicted <MAT>
 F:2-1443,3585-3924/Product: ankyrin 2, short form #status predicted <MA2>
 F:63-95/Domain: ankyrin repeat homology <AN01>
 F:96-128/Domain: ankyrin repeat homology <AN02>
 F:129-161/Domain: ankyrin repeat homology <AN03>
 F:162-190/Domain: ankyrin repeat homology <AN04>
 F:191-223/Domain: ankyrin repeat homology <AN05>
 F:232-264/Domain: ankyrin repeat homology <AN06>
 F:265-297/Domain: ankyrin repeat homology <AN07>
 F:298-330/Domain: ankyrin repeat homology <AN08>
 F:331-363/Domain: ankyrin repeat homology <AN09>
 F:364-396/Domain: ankyrin repeat homology <AN10>
 F:397-429/Domain: ankyrin repeat homology <AN11>
 F:430-462/Domain: ankyrin repeat homology <AN12>
 F:463-495/Domain: ankyrin repeat homology <AN13>
 F:496-528/Domain: ankyrin repeat homology <AN14>
 F:529-561/Domain: ankyrin repeat homology <AN15>
 F:562-594/Domain: ankyrin repeat homology <AN16>
 F:595-627/Domain: ankyrin repeat homology <AN17>
 F:628-660/Domain: ankyrin repeat homology <AN18>
 F:661-693/Domain: ankyrin repeat homology <AN19>
 F:694-726/Domain: ankyrin repeat homology <AN20>
 F:727-759/Domain: ankyrin repeat homology <AN21>
 F:760-792/Domain: ankyrin repeat homology <AN22>
 F:793-825/Domain: ankyrin repeat homology <AN23>

Query Match 6.2%; Score 550; DB 2; Length 3924;

Best Local Similarity 20.4%; Pred. No. 2,8e-23;

Matches 417; Conservative 279; Mismatches 727; Indels 626; Gaps 78;

QY 14 EENITPALALEKCDVDVERNECGTFPLMLAEGQNVBIYELLKNGANCULEDDMT 73
 Db 73 KEHGVGLVOELLGRSSVDSATRKNTALHLIASLGAQAEVVLVKEGANINAQSONGFT 132
 QY 74 ALISASKEGHINIVEELKSGAS-----LEHRDMG--- 103
 Db 133 PLYMAAENHIDVYKLLLENGANOSTATEDGFTPLAVALAQQHNGAVALLLENDKRGVR 192
 QY 104 -----GWTALMMACTGRTDVELL 123
 Db 193 LPALHTAARKDPTKSAALLLQNDHNADVQSKMNVNRTESGFTPLHIAHYGNVAVATLL 252
 QY 124 LSHGAMPSTYGLDYSTYPIIMAAGRHADIVHLLQNGAKVCSDKYGTTPYMAARG- 182
 Db 253 LNRGAADVFTA-RNGITPLHVASKRNNTMMVKLLDRGGQIDAKTRDGLTTHCAARSQH 311
 QY 183 -----HLECVKHLAMGADVDEGANSMTALTIV 210
 Db 312 DQYVELLEBERGAPLARTKNGSLPLHMAAGDHVECVKLLDHKAPVDYDTIDYLTALHY 371
 QY 211 AVKGTQSYKELIKRNPVNLTDKDGNTALIASKE----- 247
 Db 372 AAHCYRVTKLLDRANPNARALNGFTPLHACKKRIKYMELLVKYGASIOATSESG 431
 QY 248 -----GHETVODLLDAGTYNINIDRSGDVTYLGAVGSHVEIVRALLQYADID 297
 Db 432 LTPIHVAAPMGHLNIVLLLLQNGASPDVNTINGETALHMAARAGOVEVRCILRGALVD 491

```

OY 298 IRGQDKTALYMAVEKGNATWYRLQCNPDTEICTKQCEPLIKATKRNIEVELLD 357
D 492 AAREQVPLHATSLRIGKTEIVOLLQNHAPDAATNGYPLHISAREGOVDASVLE 551
OY 358 KCAKVSAYDKKQDTPHVAIRGRSRLAEILR-----NPKD 394
D 552 AGAASLAKKQFTPLHVAKAGSIDVAKLLQRRAAADSAGKNLTPHVAHNDOKV 611
OY 395 GILL-----YRPNKAGETPNIDC-SHOKSILTQI-FGAR-----HLSPT 432
D 612 ALLLEKASPHATKNGTTPHIAKKNOMQIASLTLNAGETNIVTKOGVPLHLASQ 671
OY 433 ETIDGWL-----GYDYSALADILSEPTMQPCVGLYQWGSCKFLKLEDEM-- 484
D 672 EGHITMVTLLDKGANIMHSTKSLTS-----LHLAAQ-----BDKVN 710
OY 485 -----KTPAGQTEPLFQFSWLIVFLTLCC--GGGLGV-FAFVDNLAIAISLFTAL 536
D 711 ADILTKHGADDAHFKRLGYPLIV-----ACHYGVKVNFLKGGANVNAKTKNGYPL 765
OY 537 -----IYIFVIYFGGRGEGSMMANALSTRLARHIGYLEFLKLFVNP-- 584
D 766 HOAAQOGHTHINVLQGAKNATANGNTALA--IAKRLGYSVVDTLKVVTEVYTT 823
OY 585 -----ELPEOTRALPYRFLFTDYNRLSSVGETSLAEMATLSDACEREFGLA 634
D 824 TTTTEKHKLNPETMEIVLDVSDDEGD-DIMTGCGGELRPEDUKELGDSLSOSPDL 882
OY 635 TRLEFVFTESQGGKKKKCTCCLPSEVIFLEIVGCIAGITLALFVRDPKHLTVAIL 694
D 883 GNNYLRYSLEGGSDS-----LASFSDSHSLTSASY 915
OY 695 ISIASVGLAFVL-----NCRFW--QYLDLSLNSQRRKLSASKLHL 737
D 916 LQDSAVMDSVYIPSHOVSTLAKERNSTYLSWGTENLDVAALS--SSPH-- 965
OY 738 KSEGMVYKCEVELMAEMARTIDSFONQRLVYIDGLDACEODKYLQMLDTRVLF 797
D 966 --SGFLVIFMVDAR-----GGAMRCGRHNGLR--IIPRCCTAPTRVTCGLVVRHRLAT 1016
OY 798 KGPFLAIFASDPHIITKAINONLSVLDNSINGHDYRN--IYHLVFLNSRGLSMAKF 856
D 1017 MPWVE-----GGLASRLIEVSPSGAQLGPVIVELPHFAALRG--KEEEL 1061
OY 857 LVTSATNGDI-----TCSDTGTQEDTDRVYQNSLGEWTKSGTALNRDTRRROMOR 912
D 1062 VYLRSENGDSMKHEHCQYT-----EDELINELN-GMEVLDSPDELEKRIKR 1108
OY 913 TITRQMSFDLTKLVTEWESDISPQMRLLNIVSYTGRLLRN-----QITENMDRLA 967
D 1109 IITRDF-----POYFAVVS--RIKODSNLIGEGVALSTVPOQVAVPEGALT 1156
OY 968 SWINL-TEQWYRTSMILYLEETEGLDQMTL-----KTMVERSKINPTTK--DVEPL 1019
D 1157 KRIRVGLQAOQPMHSVLAKKILGNKATSPVTLPRRRKFKPTMTITPVKASSDY-ML 1215
OY 1020 LEIDGDIRNEFV-----LSRTPVLAADVKTFLPCTVNDLPKREITAD 1065
D 1216 NGFGDADPTLRLLCSITGTPPAQMEDITGTPPLTFNECSF--TTNV--SARFPLID 1270
OY 1066 VRAAREQINIGGLAVPLPLHEGPPRPSGYSQASVSSASFNGPPGCVSPQSHSY 1125
D 1271 CROIQESVTFASQYRREI-----IC-----VPYMAK 1296
OY 1126 YSGLSGPOHPFYNNRAAVPATSSLLSMYDVYCEKRLQEGIDQMMPOYCTTIKAN 1185
D 1297 FVFAKSHDPIEAR-----LRCKMDDKDKDKTLEQ-----QENFEVARSRDVEY 1342
OY 1186 INGR-VLSQC--NIDELKEMANNGDMHLFRSMVLTMRSESVQVVEDPRLFNENSAP 1242
D 1343 LEGRPIYVDCGNIIVPLTKS-----GQHIFLFSFAFK-----ENNL--PLFYKVRDTQ 1389

```

```

OY 1243 VPHESSARRSHTEPLTEUSSQTPYTLNFSPELNTLGLDECAPRHSNI----- 1292
D 1390 EPCGR-----LSFMKPEKSTRGLVHOAINLNTLTPYTKESE 1427
OY 1293 SWSQOTRTPSLSLNSQDSSIEISKRLTDVQVQAEYRD-AVREYIAQMSQLEGTSSTIS 1351
D 1428 SDQOEELIMTISKNDDETSTETSVLKSLVNEVPYVLAISPDLISEVSEKKODLMT-- 1485
OY 1352 GRSSPHSTYYIGOSSGSGSIH-----STLEOERKREGELKOEDGRKSFMLKRGVDIYSS 1406
D 1486 -----ALITPDVSDKAGSIKVELVYKAABEEDGEPPEIYER--VKEDEKVEIL--R 1534
OY 1407 SGVSTNAPSLDP-----ITEE-----DEKSDGSKL-----LPG 1437
D 1535 SGCTRDRESSVQSRSERLGVEEWYIVSDIEEROKAPLEITEYPCVEYRIDKEING 1594
OY 1438 KKSERPSL--FOTD-----LTKLKG-----GLRYOK-----LPSDD 1468
D 1595 KVEDSTGLVNYLTDLNTCVPLPKRQLQYVQDQAKKCEALAVAGSSSEKGDTPDET 1654
OY 1469 ESGTGRVQITPHCSKMITRRLAKORECASPOEHSAPETIRPIKAKEYLSDALLDKDS 1528
D 1655 QSTOKHQPSPGKIKKPYR--FKLKEKOKKEEGLQASAE-----KAEL-----KGS 1699
OY 1529 SDGVRNENESSPHNSLHNEAADSOLE-----KANLELEDEHSGKRGMPHSL 1577
D 1700 SEESLGRDPLAEPPLPYKATSPLEETPIGSKRKVALOKRVEDE-QKGRKPLIRV 1758
OY 1578 SGLDPTLIARMSICSEBCKSPSECSLIASSPESBPACOKAYNLNTPSTVTLNNNAPT 1637
D 1759 KGED-----VPKTIHRPHAS-----PSLSERAPSPSPKTERHSTLSS 1802
OY 1638 NRANQNDLEIGIREISOVLIRGPPSP-NPTAVQNNLKSMAKRSORSYRLSKDASE 1696
D 1803 SATERHPVPSPSKTEK-----HSPVSPSA-----KTERHSPASSSKT-----E 1843
OY 1697 LHAASEST 1705
D 1844 KHSVPSPST 1852

RESULT 12
149502
ankyrin - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 13-Aug-1999
C:Accession: I49502
R:White, R.A.; Birkenmeier, C.S.; Peters, L.L.; Barker, J.E.; Lux, S.E.
Mamm. Genome 3, 281-285, 1992
A:Title: Murine erythrocyte ankyrin cDNA: Highly conserved regions of the regulatory
A:Reference number: I49502; M0ID:92345717; PMID:1386265
A:Accession: I49502
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1862 <RES>
A:Cross-references: GB:M84756; NID:9191939; PIDN:AAA37236.1; PID:9191940
C:Gene: Ank-1
C:Superfamily: ankyrin; ankyrin repeat homology
C:Keywords: alternative splicing
F:40-72/Domain: ankyrin repeat homology <AN01>
F:73-105/Domain: ankyrin repeat homology <AN02>
F:106-138/Domain: ankyrin repeat homology <AN03>
F:139-167/Domain: ankyrin repeat homology <AN04>
F:168-200/Domain: ankyrin repeat homology <AN05>
F:201-233/Domain: ankyrin repeat homology <AN06>
F:234-266/Domain: ankyrin repeat homology <AN07>
F:267-299/Domain: ankyrin repeat homology <AN08>
F:300-332/Domain: ankyrin repeat homology <AN09>
F:333-365/Domain: ankyrin repeat homology <AN10>
F:366-398/Domain: ankyrin repeat homology <AN11>
F:399-431/Domain: ankyrin repeat homology <AN12>
F:432-464/Domain: ankyrin repeat homology <AN13>

```


Db 1249 TGGTSPAQWEDI-----TGTTPLTFIKDCVSF---TTNV---SARFWLADCHQVLETVG 1296
 Qy 1075 IGGIAYPPLPLHGGPPRRPGSYGQSPASVCSASFNGPFGGVSPQPHSSYSGLSGPPH 1134
 Db 1297 LAFQLYREL-----ICV-----PYMAKEY-----VFAMKMDP-- 1323
 Qy 1135 PFYRAAVPAVPTGSSLLSMTVDVCEKLRQIEGLDQNMMPQYCTTIKKANINGRVLSDQ 1194
 Db 1324 -----VESLRNFCMTDTRKDKLTLEQGE----- 1346
 Qy 1195 NIDELKEMAMNGDWHLFRSMVLEMRVSQVVPEDPFLN-ENSSAPVHGE-----S 1248
 Db 1347 NFEEV-----ARSKDIEVLEKRPYVDQYGNLAPLTKGQOOLVFN 1386
 Qy 1249 ARRSHTLEP-----LTELSSOTPYTLNFSPFELNTGLDGBGRHSLMSQSTRRPPL 1304
 Db 1387 FYSGENKRLPPSIKIRTSQPCGRUSFLKERTTKGLPOTACONLITLPAKKTETES- 1445
 Qy 1305 SSLSNDSQSIETISKLTQDKVOAEYDAYRE---YIAQMSOLEGGTSGSTISGRSPHSTYV 1361
 Db 1446 -----DQDEIE---KTDRQSPASIALRKRYSYLTERGMTERGTAT---RSLP---TTY- 1492
 Qy 1362 ICGSSGSGSIHSTLEQERKGEKELQEDGRKSFIMKRGVITYSSGVSITNEASPLDPT 1421
 Db 1493 -----SYKPFESTPRYSQSWTAPITVPGPAKS-----GFTSLSSSSSNTSPASPLKSIW 1541
 Qy 1422 EDEKSDQSGSKLLPGKSSRPSPFOTDLKLGGLRQKLPDDEDESGTGRVQUTPHC 1481
 Db 1542 -----SVTSPPIKSTL-----GASTTSSVKSLSDV 1567
 Qy 1482 SKMIRTRRLAKARECAPQESHSAPIRTFIKANEY---LSDALLDK-----KDS 1528
 Db 1568 ASPIRSLRT-----MSSPIKTVVQSOPYNIQVSSGLAPAPVTEATPLKGL 1614
 Qy 1529 SDSCVRSNESSPNHSLINEAADSQLEKANILEDEHGSGKRMPS----- 1576
 Db 1615 ASNSTFSSRTSP-----VTTAGSLRERSITMTTPAPSPKNIMYSSSLPFXSIITNSAA 1668
 Qy 1577 -----LSGLDDPITIAMSIQSEDKKSPSECSLIASPRESMWA----- 1614
 Db 1669 PLISSPLKSVSPKSVYDISSAKITMASS---LSSPVKOMPHAEVALVNGSISPLKY 1725
 Qy 1615 CQKAYNINRTPSTVTLNNN-TAPTRANO-----NFEDEIGIRETSQYI---LRGSPN 1665
 Db 1726 ASSSTLLNGKATATLEKISSATNSVSVSAATDVEKEVFTTAMPSPSLKSYSA 1785
 Qy 1666 PTAQONENLKMMAKRBORSYTRLSKASFLHAASEST 1705
 Db 1786 PSARFOSLTPPSAS-----ALYTSL---GSSISATTSVT 1816

RESULT 14

S37771

ankyrin, erythrocyte - mouse

C:Species: Mus musculus (house mouse)

C:Date: 27-May-1994 #sequence_revision 03-Aug-1995 #text_change 13-Aug-1999

C:Accession: S37771

R:Bikemeier, C.S.; White, R.A.; Peters, L.L.; Hall, E.J.; Lux, S.E.; Barker, J.E.

J. Biol. Chem. 268, 9533-9540, 1993

A:Title: Complex patterns of sequence variation and multiple 5' and 3' ends are found at

A:Reference number: S37771; MUID:93252825; PMID:8486643

A:Accession: S37771

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1848 <BIR>

A:Cross-references: EMBL:X69063; NID:9311816; PIDN:CAA4801.1; PID:9311817

C:Superfamily: ankyrin; ankyrin repeat homology

C:Keywords: alternative splicing

F:48-80/Domain: ankyrin repeat homology <AN01>

F:81-113/Domain: ankyrin repeat homology <AN02>

F:114-146/Domain: ankyrin repeat homology <AN03>

F:147-175/Domain: ankyrin repeat homology <AN04>

F:176-208/Domain: ankyrin repeat homology <AN05>

F:209-241/Domain: ankyrin repeat homology <AN06>
 F:242-274/Domain: ankyrin repeat homology <AN07>
 F:275-307/Domain: ankyrin repeat homology <AN08>
 F:308-340/Domain: ankyrin repeat homology <AN09>
 F:341-373/Domain: ankyrin repeat homology <AN10>
 F:374-406/Domain: ankyrin repeat homology <AN11>
 F:407-439/Domain: ankyrin repeat homology <AN12>
 F:440-472/Domain: ankyrin repeat homology <AN13>
 F:473-505/Domain: ankyrin repeat homology <AN14>
 F:506-538/Domain: ankyrin repeat homology <AN15>
 F:539-571/Domain: ankyrin repeat homology <AN16>
 F:572-604/Domain: ankyrin repeat homology <AN17>
 F:605-637/Domain: ankyrin repeat homology <AN18>
 F:638-670/Domain: ankyrin repeat homology <AN19>
 F:671-703/Domain: ankyrin repeat homology <AN20>
 F:704-736/Domain: ankyrin repeat homology <AN21>
 F:737-769/Domain: ankyrin repeat homology <AN22>
 F:770-802/Domain: ankyrin repeat homology <AN23>

Query Match

5.8%; Score 516.5; DB 2; Length 1848;

Best Local Similarity 19.6%; Pred. No. 7.8e-22;

Matches 397; Conservative 281; Mismatches 651; Indels 697; Gaps 81;

Qy 21 LKALLECKDVERNECCGTPPLMAEGCNVEITVELKNGANCULEDDMTALISAK 80
 Db 98 VRELIVYGANVNAOSOKGFTPLYMAQENHLEVLVFLLENGANQNVATEDGETPLAVALQ 157
 Qy 81 EGH----- 83
 Db 158 QGHENVVAHLINYGTKGKRLPALHIAANDTRPAANVLQDDPNEDVLSKGFPLPIA 217
 Qy 84 -----THIVEELKSGASLEHRDMGWTALIMACYGRDVELLSHGANSPTGLQYS 138
 Db 218 AHYENINAVQALLNBRASVNFPPQNGITPLHLIASRGNVIVRLLDRGAQLE-TRTDE 276
 Qy 139 VPIITAGRGADIVHLLONGAKYVCSDKYGTPLVYAAKRGHLECYKHLIAGADVD 198
 Db 277 LPLHCAARGHVRISIELLDHQAIOAKTKNGLSPIHMAAGDHDLCVRLLOVNAELD 336
 Qy 199 -----QEGA-----NSMTALIVAKGYQSYKEILK 225
 Db 337 DITLDHLPLHVAHAGCHHRVAKVLLDKAKRPSRLNFTPLHIAKKNHRIWELLK 396
 Qy 226 RNPVNLTDKDGWTALIMASKEGHIEYVDLDAGTYVNIIPRSGDTVLIGAVRGHVEI 285
 Db 397 TGAIDAVTESGLTPLHVASFMGHLPIVKNLLORGASPNVSVKYEPLHMAARAGHVEY 456
 Qy 286 VRALLQYADIDIRGDNKTALYVAVEKGNATVVRILQCNPDTEICTDGETPLIKATK 345
 Db 457 AKYLLQKAKANAKADDQPLHCAARIGHGTGNVKKLLLENGASPNLATTAGHTPLTAAR 516
 Qy 346 MRNIEVEVELLDGKAYSAVDKGDTPPLVAIIGRSRIELLRPKGRRLPYPRKAG 405
 Db 517 EGHVDTALALLEEAQACMTKKGTPLHVAAYKGVRLAELLEDHDA-----HPNAG 570
 Qy 406 E---TPYNDICSHQ-----KSLI-----TQIFGAR----- 427
 Db 571 KNGITPLHVAVHNHNNIDIVLLLRGSGSPHPAMNGYTPHLIAKQONQIEVARSLDYG 630
 Qy 428 -----HLSPTFEDGMDLGYDLYSSALDIIEPTMOP-----PIC-- 462
 Db 631 SANAESVGYTPHLIAAQEGHTEMVALLSKQANGMLGNKSGTLPLHLYSQEGHVPADV 690
 Qy 463 -----VGLYQWQSGK--SPLKLKLD--EMKTFAGQOOTEPLQFQSWL 501
 Db 691 LIKHGTVVADTTMGATTPPLHVAHYGNIKIKYFLLOHODVNAKTKLG--SPLHQAAQ 748
 Qy 502 --IVFTLLLGGLGLVFAFPVD-----TNLAISLSPALVYIFVYIFGGRRGE 553
 Db 749 GHNDIVYLLKNG-----ASPNEVSSNGTTPLAIAKRLYISVTDVLAIVT----- 794
 Qy 554 SWMMANALSLARLHIGYLELLEFLKMFVNPPELPBQTTALPV---RLFTFYNRLS-- 608

```

Db 795 --DETSTVYLVSDKHMSTPEYDELDVDS---EDETSTAHISIMDELVSGKAREDSRD 848
QY 609 VGGFSLAMATLTSDEACEREFGLATLFR-----VFRE-ESQKKKKKKTKCCL 658
Db 849 VEEKEELDFVFKLDQVVE-----SPAIPRICTPPTVIRSEDOQAKSEYDESLI 902
QY 659 PEFVLEFIVGCIIGITLLAIFRVDPKHLFVNALISASVAGLAFNLNRTMVOVLDS 718
Db 903 PS-----SPATEISDNI-SPVASVHTGFLVSP-----WDA 933
QY 719 ILNSGRKLHSAASKLHLKSEGFVKLCEVELMARAKTIDSEFTQOQLRVYLIDGLD 778
Db 934 RCGSMKRSRHN-----GLRVIP-----PRCAAPTITCTCLV----- 966
QY 779 ACEQKVLQMDTVAVLFSGKGFALIFASDPHIIIIKAINQUNLVLRDSTNNGHDYMRN- 837
Db 967 -----KPOKLTTPPL-----ABEEGLASIIIALGPTGQOFLSPV 1001
QY 838 IVHLFVFL-NSRGLSNARKFLVTSATNGDITCSPTTGTQEDTRRVSONSLGEMTKLGSK 896
Db 1002 IVEIDHFASHG---DELYVLRSENGSV-----WKHKSRIGESYLDQILN-GMD 1049
QY 897 TALNRDYYRRQOMORTTQMSFDLKLVT---EDWFSDISPQTMRLNLIVSYGRL 953
Db 1050 EELGSELEKRRVCRIIT--TDPELYEVIMSRLOQD-YDTIGEG-----GSLRSKL 1099
QY 954 LRANQITFNMRLASWIMLTBOWPYRTSMILLYLEETPGLDQOMTKMYERISKNIPTT 1013
Db 1100 VPLVQATPENAIVTKVLAQ-----ADPVDELVTKLQNO----- 1137
QY 1014 KDVEPLLEIDDINFEVFLSSRPVLVARDVKTFELPCTVNLDPKLEIILADVAREQI 1073
Db 1138 -----ATFSP-ITVVEPRRRRHRPI----- 1157
QY 1074 NIGGLAYPRLPHHEGPPRPSPGYQOPAS---VCSASANGFPFGCVVSPQSHSYGSL 1130
Db 1158 ---GRIPLPPSWTDNPR-DSGEGDPTSLRLCSVI-----GTDQAO-----WEDIT 1201
QY 1131 GPOHFFY-NRAA-----VPATG-----SLLTSMTVQVCEKLRQISGLD 1170
Db 1202 GTTKLIYANECANFTTVNSARFMLSDCRTAEVHFALILYKELTA----- 1247
QY 1171 QNMMPQYCTTIKKANIN---GRVLSQCNIDE-LKKEMANNGMHLFRSVILEM----- 1220
Db 1248 ---VPYMAKFYIFAKMDAREGRRLCYCMTDKDKYLEOHEFVEVARSHDIVLEGMP 1304
QY 1221 --RSVESQVDP-----EDRPLNENSSAPVPGESARR-----S 1252
Db 1305 LFAELSGNLVYVKAQAQORSFHFQSFERNRLAIFVKVYDSSREPGLSLFLRTKMYEDT 1364
QY 1253 SH-----TELPLELS-----SOTPYLNFSEFELNLTGLDEGAPR----- 1288
Db 1365 OHILCHLNTIMPCTKSGAGDRRRITLPLRLKRSILSESLGTSTDTDEYEMRAVIRE 1424
QY 1289 HSNLSWOSQTR-----RTPSLSLINSODSSI-----EISKLDKVOAET 1327
Db 1425 HLGSLMAELAELOFSEVEDINRIYENPNLSLDOSTALLTLTMDREGENAKM-ENLYTAL 1483
QY 1328 RDAVREYIAQMSOLEGGTSGSTISGRSSPHSTYYIGQSSSGSHSTHLEDBREKGEGLKQ 1387
Db 1484 RNIRSEIVNMLE-----VSGROS-----RNLLKER----- 1509
QY 1388 EDGRKSFMLKRGDIDYSS-----SGVSTNEASPLDPIEEDKSPQSSKLLPGKSS 1441
Db 1510 RHGOREYSLSPSQVNGSSLODELSPASLOIYALP-SPLCADQYMWVYVYIDALPLAATE 1568
QY 1442 ERPSLFOTDLKLGKGLRYOKLPSEDESGTGRVQITPHCSKMIKTKRLAKOREGASPO 1501
Db 1569 HDTMLEMSDMQVMSAGLT-PSLYVTAEDSS-----LECSKAE 1603
QY 1502 EHSAPERTFIKAKEYISDALDKKDSGVSNSNESSPNHSLHNEAADSOLEKANLIE 1561
Db 1604 D-----SDAIPWK-----LEGAHSETOGPELGSDOLVE 1633

```

```

QY 1562 LEDEHSGKRMGPHSLSGLODPIIARMSICSEDKSPSECSLI-----A 1605
Db 1634 DDTVDSATNGLADLLG--QORVHARIT-----DSPSVQVOLFDRSQARTLMDKQSGTA 1685
QY 1606 SSPE-----SW--PACQAKANLNRTPSYTLNNKTPAPTRANDNPEI-----EG 1650
Db 1686 VHQEATQSSMOEBEVTQGFHSFQRITTI-----GQPEGALQETQVSVSTREHYQRP 1740
QY 1651 RETSOVLIRPSPNPPTA-----VONENLKSMAHKRSRSYST 1688
Db 1741 PETGSP--KAGKEPSLMAPEASAFQEVQDELQNTIGEDVTEQFT 1784

RESULT 15
T42714
ankyrin 3, splice form 2 - mouse
C:Species: Mus musculus (house mouse)
C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000
C:Accession: T42714
R:Peterson, L.L.; John, K.M.; Lu, F.M.; Elcher, E.M.; Higgins, A.; Yialamas, M.; Turtzo
J. Cell Biol. 130, 313-330, 1995
A:Title: Ankyrins (epithelial ankyrin), a widely distributed new member of the ankyrin ge
the repeat domain.
A:Reference number: 222237; MUID:95340633; PMID:7615634
A:Accession: T42714
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1765 <PEP>
A:Cross-references: EMBL:L40632; NID:9710548; PID:9710551; PIDN:AAB01605.1
A:Experimental source: strain C57BL/6J; kidney
A:Genetics:
A:Gene: Ankyrins
A:Map position: 10
A:Introns: 1587/1
C:Superfamily: ankyrin; ankyrin repeat homology
C:Keywords: alternative splicing

Query Match 5.88; Score 510.5; DB 2; Length 1765;
Best Local Similarity 20.18; Pred. No. 1.6e-21;
Matches 396; Conservative 278; Mismatches 641; Indels 653; Gaps 78;

QY 14 EENINPALKALEKCDVDERNECQTPMLAEOGNVEIVKELKNGANCULEDDMWT 73
Db 66 KESHEVVESELLQREANVDAAKRGKTAHISLAGOAEVAVVLYTNGANVAQSONGFT 125
QY 74 ALISAKKEGHINIVEELKSGASLEHRDNGMTALMMACYKRTVVELL----- 124
Db 126 PLYMAOENHLEVVRLDNGASQSLATEDGFTPLVALQGHQDVVSLLENDTKGKVR 185
QY 125 -----SHGANPSYTGK----- 135
Db 186 LPLALTAARKDDTKAALLLNDNTYADVESKSGFTPLHIAHYGNINATLLNRAAAD 245
QY 136 ---QYSVYPIIWAAGRHADIVHLLONGAKVNCSDKGTPTLVAAARK----- 182
Db 246 FTAANDITPLHNASKRGANNMVKLLDRAKIKDATIRGILPLHCGASGHEQVYEMLLD 305
QY 183 -----HLECYKHLIAMDADVDEGANSMTALIVAAGKYQTS 219
Db 306 RSAPILSKTKNGLSPLHMAQGDHLCVQLLQHNVPDYVDYLTALHYAAGCHYKV 365
QY 220 VKEILLKRNNVNLTLDKGTALAMISKE----- 247
Db 366 ARVLLDKKASPAKALNGFTPLHACKNRIRIYRMELLKHGASIOAVTESGLTPIHYA 425
QY 248 -GHIEVODLDAGTYVNIPIRSGDVLIGAVRGHVEITVALLQKYADIDRQDNKTA 306
Db 426 MCHVNIYVSLMHNHGASPTTNYRGTALHMAARSQAQAVVYVYVODGQVAKAKDDQTP 485
QY 307 LYMAVEKGNATVNRDILQCNPDTEICTYKDEGTPLIKATKMRNIEVELLDGAKVSAVD 366
Db 486 LHISARLGRADIVQOLLQOGASPNNAATTSGYTPLHILAREGHEBVAFLLDHGAASLSTTT 545

```

```

QY 367 KKQDTPLHVAIRGRSRLAELLR-----NPKDGRLL----- 398
Db 546 KKQDTPLHVAIRGRSRLAELLR-----NPKDGRLL----- 398
QY 399 --VRPNKAGEPYNIIDC-SHOKSILNDI-FGAR-----HLSPTETDGMGLY 441
Db 606 SPHAAKNGYTPPLHIAKKNMOMIAISLEYGADANAVTROGLASVHLAQBEGHVDVSL 665
QY 442 LYSXALADILSEPTMOP-----PIGVGLYAOV 469
Db 666 LLSHNAVNLSNKGSLPPLHIAQEDRVNAEVLNCGAHVDAQTKMGYPLHVGCH--Y 723
QY 470 GSGK--SFLK--KLEDEKKTFRAGQOTEPLFOESWILYFLT-----LLCGGLGV 516
Db 724 GNKIVFPLLOHSAKNAKTKNGYTAHQAOQHHTHINVLONNASPNELTVNG---- 779
QY 517 FAFVDTNLAIASISFLALYIFFIY----- 544
Db 780 -----NTALAIARLGIVSVVDLKYVTEELMTTTEKHKNVPEPTNMEVLDMSDEV 834
QY 545 -----YFGRRGESWMMALSTRLAHIGYLELLEKL-----MFIN 562
Db 835 RKASAPKLSDEGEYSIDSEEDCKTWPFIKPVQ-----EVLKSEDAITGDDTYLG 886
QY 583 PRELEQTTALPVRFLFTDYNRLSVGETSLAEMITLSDACEREFGLATPLRFR 642
Db 887 PDLKEIGDSDLPAR-----GYVG--FSLGARSASLR-----SFSSDRSYTLNR 928
QY 643 TEESQGRKKWKTCCLPSFVIFLFTVGCIIAGITLAIIFRYP-KHLY-----VNAI 693
Db 929 SSYAR-----DSMIEELLPVSKBQHLLFTREPOSDSLRHSMAADTLDNVNLV 977
QY 694 LISTASYVGLAFVINCRTMVOVLDLSNQ-----RK-----RLHSASKLKL- 737
Db 978 SSPVHSGFLVSVFVNDAG-----GSMGRSRRHGMRIIPRKCCTAPRTICRLVRRKLA 1032
QY 738 -----KSEB-----FMKVLKCEVELMARMAKTIDSFON 766
Db 1033 NPPWVEBEGSLASRLVEMKPRGAOFLGPVIVEIFHFGSMGRKRELIVLSSENETYKH 1092
QY 767 Q-----TRLVYIDGLDACEQDKVLOMLDLYRV--LFSKGFIAIFASDPHIIKAINON 819
Db 1093 QEDSKNEDLAEILLNGMD--EELDSPEELGTFRICRIITK-DEPOYFA-----VYSRIKOE 1144
QY 820 LNSVLRDSNNGHDMRYIHALPVLNSRGLSNARKFLV--TSATNGDITCSDTGTQ--- 875
Db 1145 SNOIGPEGI-----LSTTVPLVOASPEEGALTKRIRVGLQAOP 1184
QY 876 --EDIDRRVSONSLGEMTKLSKTALNRRDY--RRROMRTITROMSFDLTKL--VT 928
Db 1185 VPEELVYKI-----LGNKATFSPVIVEPERRRKFHPT--MTIIPPPSGEGYS 1232
QY 929 EDWFSDISPQTRRLINIVSYTG-----RLRANOITENMDRLASWINLTBQWRYT 980
Db 1233 NGYKGDATPML--RLL--CSITGTSPOWEMDITGTPLFRIKDCVSFTVNSARF--- 1284
QY 981 SWL-----ILYLEBEGLPDOMTKMTERISKNIPTTKDVEPLLEIDGDIRNFEVPLSS 1035
Db 1285 -WLADCHQVL--ETVGLASQLYRELIC-----VPR-----MAKFVFAKT 1321
QY 1036 RTPVLVARDVKTFLPCTVNLDPKLEITADYARAREOINIGLA-YPLPLHEGPPRPPS 1094
Db 1322 NDP-----VSSLKRCFCMTDRDKTL-----EQENFEVARSKDIEVLEBKPIYVD 1369
QY 1095 GYSQAPAVCSSASFNGPPGVSPQSHSYSGISGQHPF--YNRAAVPATGSSLL 1150
Db 1370 CYGNALPLTK-----GG--QOLVFNEYS--FKENRLPFSIKIRDTSOEPCGRISFL 1416
QY 1151 LSSMTVDVCEKRLQIEGLDNMMPOYCTTIKKANINGRVLSQCNIDELKEMAMNFGDW 1210
Db 1417 KEPKTT-----KGLPOTAV-----CNLNTLPLAHKKAERKDRQS----- 1451

```

```

QY 1211 HLFPSVULMR--SVESQVPEDEPFLNENSSAPVPHGESARSSHTELPTELISQTP 1267
Db 1452 --FASIALKKRYSYLETPSMSPQSPCERDIMALY-----ADHGLSTELARE-- 1499
QY 1268 YTLNFSFEELNTLGLDEGAPRHSNL-----W--QOQTRTPSLSSLNSQDSSTET 1316
Db 1500 --LNFVDEILNQIRVEN--PNSLISQSPMLKKWYTRDGNKATDALTSVLTKINRIDI 1554
QY 1317 SKLTDKVOAEYRDATREYTAQMSQLEGGGSSITSGRSSPHSTYIG--QSSSGSIHST 1374
Db 1555 VTLLEGPFIPIYDNI-----GTRSFADENNVEFDPVDGMQNETPSGSLSP 1600
QY 1375 LEOERKGEDELKOEODRKSFLMKRGVDIDYSSSGVSTNEASPLDPTTEDEKSDQSG-- 1431
Db 1601 AQARRLTGLDLRLD-----DSSQARDSTSY-----LTGEPKIEANGNHT 1643
QY 1432 SKLLPKKSSERSPLFOTDLKLGGLRYQKLPSDEDESGTGRVOITPHCSKMITRK 1491
Db 1644 AEVIPEAKAK-----PYEPESONDIGKOSIK-----ENLK 1673
QY 1492 AKQREKASPOEHSAPERTIFIKAKELYSDALD--KKOSSDSGVRSNESSPNHSLHNEA 1549
Db 1674 PKTHGCGRTBE-PVSPLTAYOKSLEETSKLVIEDAKPCVPVGMK-----KMTRTTA 1724
QY 1550 DDSQLEKANLIELEDEGHSGKRGMPHSLSLGLDPTIARMSICSEDKS 1597
Db 1725 DG-----KARLNLOEEGST--RSEPKQGGYK--VTKKEINRVEKKT 1764

```

Search completed: July 1, 2003, 14:42:36
 Job time: 71 secs



•
•
•
•

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2003, 14:37:09 ; Search time 38 Seconds

(without alignments)
4948.535 Million cell updates/sec

Title: US-10-021-571-2

Perfect score: 8853
Sequence: 1 MEVLISQSYINVERENIPA.....ELHAASSTGCGEERSTL 1715

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 424699 seqs, 109646833 residues

Total number of hits satisfying chosen parameters: 424699

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Published_Applications_AA.*

1: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubppa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB.pep.*
5: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep.*
6: /cgn2_6/ptodata/1/pubppa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubppa/US09_NEW_PUB.pep.*
9: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubppa/US10_NEW_PUB.pep.*
11: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubppa/US60_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2961.5	33.5	705	US-10-106-698-6378	Sequence 6378, App
2	2782.5	31.4	551	US-09-835-788A-17	Sequence 17, Appl
3	2260	25.5	513	US-10-149-819-9	Sequence 9, Appl
4	549	6.2	109	US-09-986-480-371	Sequence 371, App
5	538.5	6.1	1724	US-09-964-899-43	Sequence 43, Appl
6	438	4.9	740	US-09-835-788A-12	Sequence 12, Appl
7	401.5	4.5	426	US-09-908-711-70	Sequence 70, Appl
8	395.5	4.5	1327	US-09-972-115A-8	Sequence 8, Appl
9	395.5	4.5	1327	US-09-841-835-2	Sequence 18, Appl
10	389.5	4.4	285	US-09-835-788A-18	Sequence 10, Appl
11	387.5	4.4	949	US-09-841-835-10	Sequence 6, Appl
12	380.5	4.3	1166	US-09-972-115A-6	Sequence 15, Appl
13	380.5	4.3	1166	US-10-163-587A-15	Sequence 2, Appl
14	371.5	4.2	1074	US-09-509-196A-2	Sequence 7, Appl
15	360.5	4.1	1333	US-09-972-115A-2	Sequence 13, Appl
16	360	4.1	784	US-10-164-080-7	Sequence 2, Appl
17	355	4.0	251	US-09-835-788A-13	Sequence 2, Appl
18	355	4.0	786	US-10-164-080-2	Sequence 2, Appl
19	355	4.0	786	US-10-299-327-2	Sequence 2, Appl

20	355	4.0	787	US-09-866-050A-334	Sequence 334, App
21	354.5	4.0	1267	US-09-972-115A-4	Sequence 4, Appl
22	354	4.0	802	US-09-964-899-41	Sequence 41, Appl
23	349.5	3.9	679	US-10-339-936-2	Sequence 2, Appl
24	331	3.7	835	US-09-947-199-2	Sequence 2, Appl
25	331.5	3.6	673	US-09-841-835-8	Sequence 8, Appl
26	303	3.4	835	US-09-947-199-8	Sequence 19, Appl
27	299	3.4	306	US-09-835-788A-19	Sequence 378, App
28	278	3.1	1719	US-10-012-896-378	Sequence 378, App
29	278	3.1	1719	US-09-895-814-378	Sequence 378, App
30	278	3.1	1719	US-09-895-814-378	Sequence 378, App
31	278	3.1	1719	US-10-010-940-378	Sequence 378, App
32	278	3.1	1719	US-09-759-143-378	Sequence 378, App
33	278	3.1	1719	US-09-780-669-378	Sequence 378, App
34	278	3.1	1719	US-09-822-827-378	Sequence 378, App
35	263.5	3.0	599	US-09-735-368-2	Sequence 2, Appl
36	249	2.8	622	US-10-197-666A-22	Sequence 22, Appl
37	242	2.7	1054	US-09-798-042-87	Sequence 87, Appl
38	240.5	2.7	622	US-10-197-666A-6	Sequence 6, Appl
39	238.5	2.7	329	US-09-880-192-62	Sequence 62, Appl
40	238.5	2.7	329	US-09-758-593A-11	Sequence 11, Appl
41	236.5	2.7	328	US-09-758-593A-11	Sequence 12, Appl
42	236.5	2.7	328	US-09-758-593A-12	Sequence 1, Appl
43	236	2.7	319	US-10-312-064-1	Sequence 167, App
44	233.5	2.6	210	US-10-127-032-167	Sequence 66, Appl
45	233.5	2.6	460	US-09-905-673-66	

ALIGNMENTS

RESULT 1	US-10-106-698-6378	Application US/10106698
Sequence 6378, Application US/10106698	Publication No. US20030109690A1	
GENERAL INFORMATION:		
APPLICANT: Ruben et al.		
TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypept		
FILE REFERENCE: PA005P1		
CURRENT APPLICATION NUMBER: US/10/106,698		
PRIOR APPLICATION NUMBER: PCT/US00/26524		
PRIOR FILING DATE: 2000-09-28		
PRIOR APPLICATION NUMBER: US 60/157,137		
PRIOR FILING DATE: 1999-09-29		
PRIOR APPLICATION NUMBER: US 60/163,280		
PRIOR FILING DATE: 1998-11-03		
NUMBER OF SEQ ID NOS: 8564		
SOFTWARE: PatentIn Ver. 3.0		
SEQ ID NO 6378		
LENGTH: 705		
TYPE: PRT		
ORGANISM: Homo sapiens		
FEATURE:		
NAME/KEY: MISC_FEATURE		
LOCATION: (244)		
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids		
NAME/KEY: MISC_FEATURE		
LOCATION: (337)		
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids		
US-10-106-698-6378		
Query Match	33.5%;	Score 2961.5; DB 9; Length 705;
Best Local Similarity	95.5%;	Pred. No. 8.3e-181;
Matches	567;	Conservative 12; Mismatches 14; Indels 1; Gaps 1;
QY	62	ANCNEDLDNWTALISASKEGHIHYEELKSGASLEHRDGGMTALMACKGRTDYE 121
DB	28	SNCLNDLDNWTALISASKEGHIHYEELKSGASLEHRDGGMTALMACKGRTDYE 87
QY	122	LLISHGANSYVGLGVYPTIWAAGRGHADIYHLLONGAVNCSDKYGTPLVMAARK 181
DB	88	LLISHGANSYVGLGVYPTIWAAGRGHADIYHLLONGAVNCSDKYGTPLVMAARK 146

QY 182 GHECVKHLAMGADVDEGANSMTALIVAKGYTQSYKEITLKRNPVNLTDKGNATL 241
 DB 147 GHECVKHLAMGADVDEGANSMTALIVAKGYTQSYKEITLKRNPVNLTDKGNATL 206
 QY 242 MASKEGHEIYODLLDAGTYVNIIPDRSGDYVLIGAVGHEIYRALLQKXADIDINGQ 301
 DB 207 MASKEGHEIYODLLDAGTYVNIIPDRSGDYVLIGAVGHEIYRALLQKXADIDINGQ 266
 QY 302 DKRTALYMAVEKGNATVWDILQCNPDTEICTKDETEPLIKATKRNIEVEVELLDKAK 361
 DB 267 DKRTALYMAVEKGNATVWDILQCNPDTEICTKDETEPLIKATKRNIEVEVELLDKAK 326
 QY 362 VSAVADKKDPTLHVAIRGSRRLAELLRNPKDGRLLYRPNKAGETPYNIDCSHOKSLYT 421
 DB 327 VSAVADKKDPTLHVAIRGSRRLAELLRNPKDGRLLYRPNKAGETPYNIDCSHOKSLYT 386
 QY 422 QIFGARHLSPETEDDMIGDYLSAALDIISEPTMOPPCVGLYAQWGSKSPFLKLE 481
 DB 387 QIFGARHLSPETEDDMIGDYLSAALDIISEPTMOPPCVGLYAQWGSKSPFLKLE 446
 QY 482 DEMKTFAGQOTPELPFQFSMLYVFLTLGCGIGIYFAFPVDNTALATISFLALITYEF 541
 DB 447 DEMKTFAGQOTPELPFQFSMLYVFLTLGCGIGIYFAFPVDNTALATISFLALITYEF 506
 QY 542 IYIYFGRRGESSNMAMALSTRLARHIGYLELLFKLMFVNPPELPBQTTKALPVRFLETT 601
 DB 507 IYIYFGRRGESSNMAMALSTRLARHIGYLELLFKLMFVNPPELPBQTTKALPVRFLETT 566
 QY 602 DYNRLSSVGETSLAEMITATLSDACEREGFLATRLRFVFRTEBSQKKKKKKKT 655
 DB 567 DYNRLSSVGETSLAEMITATLSDACEREGFLATRLRFVFRTEBSQKKKKKKKS 620

RESULT 2

US-09-835-788A-17
 ; Sequence 17, Application US/09835788A
 ; Patent No. US20020077458A1
 ; GENERAL INFORMATION:
 ; APPLICANT: NI et al.
 ; TITLE OF INVENTION: Death Domain-Containing Receptor Polynucleotides, Polypeptides,
 ; FILE REFERENCE: P018P1
 ; CURRENT APPLICATION NUMBER: US/09/835,788A
 ; PRIOR FILING DATE: 2001-04-17
 ; PRIOR APPLICATION NUMBER: PCT/US00/28666
 ; PRIOR FILING DATE: 2000-10-17
 ; PRIOR APPLICATION NUMBER: 60/159,585
 ; PRIOR FILING DATE: 1999-10-18
 ; PRIOR APPLICATION NUMBER: 60/167,246
 ; PRIOR FILING DATE: 1999-11-24
 ; NUMBER OF SEQ ID NOS: 24
 ; SOFTWARE: Patent In Ver. 2.0
 ; SEQ ID NO 17
 ; LENGTH: 551
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-835-788A-17

Query Match 31.4%; Score 2782.5; DB 10; Length 551.
 Best Local Similarity 96.7%; Pred. No. 1.6e-169;
 Matches 534; Conservative 8; Mismatches 9; Indels 1; Gaps 1;

QY 102 MGMTALMMACYKRTDVVELLSHGPNPVTGLQYVSYPYLIMAGRGHADIYHLLQNG 161
 DB 1 MGMTALMMACYKRTDVVELLSHGPNPVTGL-YEYVYLIMAGRGHADIYHLLQNG 59
 QY 162 AKVNCSDKGYTTPLYMAARKGHECVKHLAMGADVDEGANSMTALIVAKGYTQSYK 221
 DB 60 AKVNCSDKGYTTPLYMAARKGHECVKHLAMGADVDEGANSMTALIVAKGYTQSYK 119
 QY 222 EILKRNPNVNLTDKGNATLMAKSGHEIYODLLDAGTYVNIIPDRSGDYVLIGAVRG 281

DB 120 EILKRNPNVNLTDKGNATLMAKSGHEIYODLLDAGTYVNIIPDRSGDYVLIGAVRG 179
 QY 282 HVEIVRALLQKXADIDIRGODNKTALYMAVEKGNATVWDILQCNPDTEICTKDETEPLI 341
 DB 180 HVEIVRALLQKXADIDIRGODNKTALYMAVEKGNATVWDILQCNPDTEICTKDETEPLI 239
 QY 342 KATKRNIEVEVELLDKAKVSAVDKKDPTLHVAIRGSRRLAELLRNPKDGRLLYRPN 401
 DB 240 KATKRNIEVEVELLDKAKVSAVDKKDPTLHVAIRGSRRLAELLRNPKDGRLLYRPN 299
 QY 402 NKAGETPYNIDCSHOKSLITQIFGARHLSPETEDDMIGDYLSAALDIISEPTMOPPCV 461
 DB 300 NKAGETPYNIDCSHOKSLITQIFGARHLSPETEDDMIGDYLSAALDIISEPTMOPPCV 359
 QY 462 CVGLYAQWGSKSPFLKLEDEKKTAGQOTPELPFQFSMLYVFLTLGCGIGIYFAFPV 521
 DB 360 CVGLYAQWGSKSPFLKLEDEKKTAGQOTPELPFQFSMLYVFLTLGCGIGIYFAFPV 419
 QY 522 DTNLAISLSPALITYEFIVYFGRRGESSNMAMALSTRLARHIGYLELLFKLMFV 581
 DB 420 HPMGLIATVSLFLALITYEFIVYFGRRGESSNMAMALSTRLARHIGYLELLFKLMFV 479
 QY 582 NPPELPBQTTKALPVRFLETTDYNRLSSVGETSLAEMITATLSDACEREGFLATRLRFV 641
 DB 480 NPPELPBQTTKALPVRFLETTDYNRLSSVGETSLAEMITATLSDACEREGFLATRLRFV 539
 QY 642 RTEBSQKKKKKK 653
 DB 540 KTEDTQKKKKK 551

RESULT 3

US-10-149-819-9
 ; Sequence 9, Application US/10149819
 ; Publication No. US20030044913A1
 ; GENERAL INFORMATION:
 ; APPLICANT: INCYTE GENOMICS, INC.
 ; APPLICANT: YUE, Henry
 ; APPLICANT: AZIMZAI, Yalda
 ; APPLICANT: TANG, Y. Tom
 ; APPLICANT: PATTERSON, Chandra
 ; APPLICANT: BAUGHN, Mariah R.
 ; APPLICANT: LU, Dying Aina M.
 ; APPLICANT: SHAH, Purvi
 ; APPLICANT: TAL, Preethi
 ; APPLICANT: AU-YOUNG, Janice
 ; APPLICANT: BURFORD, Neil
 ; TITLE OF INVENTION: EXTRACELLULAR MATRIX AND CELL ADHESION MOLECULES
 ; FILE REFERENCE: PF-0760 PCT
 ; CURRENT APPLICATION NUMBER: US/10/149,819
 ; PRIOR FILING DATE: 2002-06-10
 ; PRIOR APPLICATION NUMBER: 60/172,852; 60/172,354
 ; PRIOR FILING DATE: 1999-12-10; 1999-12-16
 ; NUMBER OF SEQ ID NOS: 42
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 9
 ; LENGTH: 513
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; NAME/KEY: misc_feature
 ; OTHER INFORMATION: Incyte ID No. US20030044913A1 2215706CD1
 ; US-10-149-819-9

Query Match 25.5%; Score 2260; DB 9; Length 513;
 Best Local Similarity 86.6%; Pred. No. 3.3e-136;
 Matches 445; Conservative 21; Mismatches 46; Indels 2; Gaps 2;

QY 1203 MAMNFGDWLFRSNTYLEMNSVESQVVPEDPRLNENSSAPVPHGSRARSSHTLEPTEL 1262
 DB 1 MAMNFGDWLFRSNTYLEMNSVESQVVPEDPRLNENSSAPVPHGSRARSSHTLEPTEL 60
 QY 1263 SSQPTTYLNFSEELNLGLDEGADPRHSNLSWQSOTRRTPTSLSLNSODSSIEISKLTDK 1322

```

Db      61 SSOTPTLNFSEELNTGLDGCAPRHSLMSQOTRRPISLINSQSSSIEIKLMDK 120
      1323 VOAEKDAYREYIAQNSQLEGGTSGSTISGRSSPHSTYIIGQSSGGSHSTLEDRKE 1382
      121 VOAEYKDAYREYIAQNSQLEGGTSGSTISGRSSPHSTYIIGQSSGGSHSTLEDRKE 180
      1383 GELKODGKSPFLMKRGVDIVSSSGVSTNEMASPLDPTTEDEKSPQSSKLLPKKSE 1442
      181 SEKPDGKSPFLMKRGVDIVSSSGVSTNEMASPLDPTTEDEKSPQSSKLLPKKSE 240
      1443 RPSLFTQDLKLGKGLRYOKLPSDEDESGTGRVQJTPHCSKMIPTKRLAKOREGASPOE 1502
      241 RSLFTQDLKLGKGLRYOKLPSDEDESGTGRVQJTPHCSKMIPTKRLAKOREGASPOE 299
      1503 HSAEPIRTFKAEYLSDALDKDSSDGVSNSSPHSLHNEAADDQLEKANLIEL 1562
      300 HSAEPIRTFKAEYLSDALDKDSSDGVSNSSPHSLHNEAADDQLEKANLIEL 359
      1563 EDGSHGKSPFLMKRGVDIVSSSGVSTNEMASPLDPTTEDEKSPQSSKLLPKKSE 1622
      360 EDGSHGKSPFLMKRGVDIVSSSGVSTNEMASPLDPTTEDEKSPQSSKLLPKKSE 419
      1623 RPSSTVTLNNMTAPTRANQNFDEIGIRETSQVILRPGSPMPPTAVQNMNLKSMHRS 1682
      420 RPSSTVTLNNMTAPTRANQNFDEIGIRETSQVILRPGSPMPPTAVQNMNLKSMHRS 479
      1683 QRSSTYRLSKDASELH-AASSESTGGERESTL 1715
      480 QRSSTYRLSKDASELH-AASSESTGGERESTL 513

```

RESULT 4

```

US-09-986-480-371
; Sequence 371, Application US/09986480
; Publication No. US20030027999A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 143 Human Secreted Proteins
; FILE REFERENCE: P5500P1
; CURRENT APPLICATION NUMBER: US/09/986,480
; CURRENT FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: PCT/US00/12788
; PRIOR FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: US 60/134,068
; PRIOR FILING DATE: 1999-05-13
; NUMBER OF SEQ ID NOS: 456
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 371
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-986-480-371

```

```

Query Match      6.2%; Score 549; DB 9; Length 109;
Best Local Similarity 98.1%; Pred. No. 6.3e-28;
Matches 105; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

Db      302 DNKTALYWAWEKGNATVVRDILQCNPTETICTKDEGTPPLIKATKMRNIEVELLDKGA 361
      1 DNKTALYWAWEKGNATVVRDILQCNPTETICTKDEGTPPLIKATKMRNIEVELLDKGA 60
      362 VSAADKKGDPPLHVAIGRSRLAELLRNPKDGRLLYRNKAGETP 408
      61 VSAADKKGDPPLHVAIGRSRLAELLRNPKDGRLLYRNKAGETP 107

```

RESULT 5

```

US-09-964-899-43
; Sequence 43, Application US/09964899
; Patent No. US20020174446A1
; GENERAL INFORMATION:
; APPLICANT: Cohen, Dalia et al.

```

```

; TITLE OF INVENTION: Identification of Genes Involved in
; FILE REFERENCE: 4-31612 A
; CURRENT APPLICATION NUMBER: US/09/964,899
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/236,893
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/298,309
; PRIOR FILING DATE: 2001-06-14
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 1724
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-964-899-43

```

```

Query Match      6.1%; Score 538.5; DB 9; Length 1724;
Best Local Similarity 20.3%; Pred. No. 2e-25;
Matches 390; Conservative 250; Mismatches 661; Indels 621; Gaps 72;

```

```

      14 EEENPALKALLEKCKVDNERNECGPTPLMLAEGNGVEIKELKNGANCLEDDNWT 73
      45 KGHAEVVSSELOREANVDAATKKGNTALHLSAGAEVVKVLTNGANVNAOSQNGFT 104
      74 ALISAKSGHIIHVEELIKSGASLEHRDMGWTALMAACYGRDVEELL----- 124
      105 PLYMAOENHELEVAFILDNAGSOSLATEDEFTPLAVALQGHQDVSLLENDTKGVK 164
      125 -----SHGANPSTYGL----- 135
      165 LPALHIAARKDTKAALLLQNDNNADVESKSGFTPLHIAHYGINVATLLNRAA 224
      136 ---QSYVPIIWAARGHADIYHLLONGAKVNSDXYGTPPLVWAARKG----- 182
      225 FTARNDDTPLHVAISRGANNAVYKLLDNKAKIDAKTRGTLPLHCGAASGHEQVEMLLD 284
      183 -----HLECYKHLAMGADYDOBANSMALIVAAGKYQS 219
      285 RAAPILSKTKNGLSPLHATQGDHLCYQLLQHNVPVDDVTNDVLTALHVAHGHYKV 344
      220 VKEILKRNPNVNLTKDGNLTALMASKSGHIEIVODLDAGTYVPIPRSGDYILGAVR 279
      345 AKVLLDKRANPAKALNGFTPLHIAKKNRIKIVMELIKHGSIDAVTEREETALHMAAR 404
      280 GGHVEYVRLALQKYADIDIRGDNKNTALYWAWEKGNATVVRDILQCNPTETICTKDEGTP 339
      405 SGQAEVVRVLYVDGAEVAKKARDOTPLHISARLGRADIVDOLLQOGASPNNAITTSGLTP 464
      340 LIKATKMRNIEVELLDKGAARVSAVKKGDPPLHVAIGRSRLAELLRNPKDGRLLY 399
      465 LHSAREGHEDVAAPLLDHGASLSITTKSGLPPLHVAHYDQKVALLL---DQASPH 521
      400 RPNKAGETPNIDC-SHOKSLTQI-FCAR-----HLSPIETDGMGLDYLY 444
      522 AAKKNQYVTLHIAAKKNQMDIATTLLEYGADANNAVYTRQGIASVHLAAGHVDWVSLIG 581
      445 SSALADILSEPMPQPIGVGLYAOWSGSKSFLIKLEDEMK-----TFAGQOTE-- 493
      582 RANVNVLSKSGILTP---LHLAQ-----EDRVNVAEVLVNGAIVDQOTVY 626
      494 -----PLQFSWL---IVFLTLGLGGLVLPAPVDTNLIAISL 531
      627 GPLPBGKCEVHLVTANGYTPPLHQAQOGHTHIIIVLLQNN---ASPNELTIVYTEKH 681
      532 SFALALYIFVIYFEGRGDEGSNNAMALSTRIRAHIGIYELLFKIMVYN----- 583
      682 KM-----NVPEPMNEVLDMSDDEVKKANAPMLSDGETISVEEGNCT 725
      584 ---PELPQOTKALPVREFLTDYNNLSVSGETSLA-----EMATLSD---ACE 627
      726 WYKIPKVOEFTYKTFKREAPDVGLISTASAGEDAMTGTDTKYLGQDLKELGDDSLPAE 785

```



```

P101001.1  PRIOR APPLICATION NUMBER: 09/764,868
P101001.1  PRIOR FILING DATE: 2001-01-17
P101001.1  PRIOR APPLICATION NUMBER: US01/01312
P101001.1  PRIOR FILING DATE: 2001-01-17
P101001.1  PRIOR APPLICATION NUMBER: 60/179,065
P101001.1  PRIOR FILING DATE: 2000-01-31
P101001.1  PRIOR APPLICATION NUMBER: 60/180,628
P101001.1  PRIOR FILING DATE: 2000-02-04
P101001.1  PRIOR APPLICATION NUMBER: 60/209,467
P101001.1  PRIOR FILING DATE: 2000-06-07
P101001.1  NUMBER OF SEQ ID NOS: 167
P101001.1  SOFTWARE: PatentIn Ver. 2.0
P101001.1  SEQ ID NO 70
P101001.1  LENGTH: 426
P101001.1  TYPE: PRT
P101001.1  ORGANISM: Homo sapiens
P101001.1  FEATURE:
P101001.1      NAME/KEY: SITE
P101001.1      LOCATION: (148)
P101001.1      OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
P101001.1      NAME/KEY: SITE
P101001.1      LOCATION: (167)
P101001.1      OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
P101001.1      NAME/KEY: SITE
P101001.1      LOCATION: (169)
P101001.1      OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
P101001.1      NAME/KEY: SITE
P101001.1      LOCATION: (258)
P101001.1      OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
P101001.1      NAME/KEY: SITE
P101001.1      LOCATION: (396)
P101001.1      OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
P101001.1      NAME/KEY: SITE
P101001.1      LOCATION: (413)
P101001.1      OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
P101001.1      NAME/KEY: SITE
P101001.1      LOCATION: (414)
P101001.1      OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
P101001.1      NAME/KEY: SITE
P101001.1      LOCATION: (417)
P101001.1      OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
P101001.1      NAME/KEY: SITE
P101001.1      LOCATION: (418)
P101001.1      OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
P101001.1      NAME/KEY: SITE
P101001.1      LOCATION: (421)
P101001.1      OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
P101001.1  US-09-908-711-70
Query Match      4.5%; Score 401.5; DB 10; Length 426;
Best Local Similarity 29.2%; Pred. No. 1.3e-17;
Matches 107; Conservative 77; Mismatches 143; Indels 39; Gaps 4
QY      26  EKCKDVERNECGQPTPLMAAGQVNEIVYKELKNGKANGCNLEDDLNMTALISASKEGHN 85
DB      30  KRSRDKRKKSADNAAYLR--AARAGHLEKALDVIKNGVDINICNONGNAILHLASKEGHVE 88
QY      86  IVEELLKSGASLEHRMGWFTPLMAYCKRGRTDVEILLSHGAPSTVGLQYSVPIVA 145
DB      89  VSEELLQREANDVAATKRGNTALHLASLAGQAEVVKVLTVGAVNNQOS--ONGFTPLYMA 147
QY      146  AGRGHADIVILLQNGAKVNCSDKGYTTPPLWAAKRGHLECVKHLAMGADVQEGANSM 205
DB      148  XQENHLEVYKFLDNGASQXKTEDEGFTPLVALQGGHDQVSVLL-----ENFTKGVRL 203
QY      206  TALIVAVKGGYQSVEELIKRNPNNVNLDDKGNALMTASKEGHIIEYODLLDAGTYNI 265
DB      204  PALHIAARKDQTKAALLQNNNDVYESKSGFTPLHATGNI----- 248
QY      266  PDNRSGDTVLLGAVRGHVEIVRALLOKVDIDIRGQDKTKLTVAAVEKGNATVRLTQC 325
DB      249  -----NVAITLLRKRAAVDFTARNDITPLHVAASKRGANMYKLLIDR 250

```

QY	326	NPTEICTDGEPLIKATMRNIEVVELLDKGAASAVDKGPTPIVAIRBSRRLA	385
	:	: : :	:
	:	: :	:
	:	: : :	:
	:	: : :	:
	:	: : :	:
Dd	291	GAKIDAKTRDGLPLHCGARSGHEQVEYLDRAPILSKTKNGLSPLHMATOGDLNCV	350
QY	386	ELLRN	391
	:	: : :	:
	:	: : :	:
Dd	351	QLLEQH	356

RESULT 8

```

? Sequence 8, Application US/09972115A
? Publication No. US20030032766A1
?
? GENERAL INFORMATION:
?
? APPLICANT: Genon Corporation
?
? APPLICANT: Gregg, Morin B.
?
? APPLICANT: Walter, Funk D.
?
? APPLICANT: Mieczyslaw, Piatyszek A.
?
? TITLE OF INVENTION: A Second Mammalian Tel
?
? FILE REFERENCE: 080/003C
?
? CURRENT APPLICATION NUMBER: US/09/972.115A
?
? CURRENT FILING DATE: 2001-10-05
?
? PRIOR APPLICATION NUMBER: US 60/128, 577
?
? PRIOR FILING DATE: 2000-04-10
?
? PRIOR APPLICATION NUMBER: US 60/129, 123
?
? PRIOR FILING DATE: 1999-04-13
?
? NUMBER OF SEQ ID NOS: 64
?
? SOFTWARE: PatentIn version 3.1
?
? SEQ ID NO 8
?
? LENGTH: 1327
?
? TYPE: PR1
?
? ORGANISM: Homo sapiens
?
? US-09-972-115A-8

```

Query Match	4.5%;	Score 395.5;	DB 9;	Length 1327;
Best Local Similarity	23.3%;	Pred. No. 1.8e-16;		
Matches 167;	Conservative 80;	Mismatches 211;	Indels 259;	Gaps 17

```

OY 375 -----VAIRGRSRELAETLRNPEDGILRYRPAKAEPT 408
Db 884 DIAALLIKYNTCVNATNDKWAFTPRHNAQKQGT-GLCALLLHAGNDPTM---KNQEQGT 939
OY 409 YN-----IDCSHOKSILT-----QIFGARHSPTEDODMIGLYVYSSALADI 451
Db 940 LDLATADDDRALLIDAMPPEALPTCKFOATVYVSSALISPASTS-----CLSAASIDN 994
OY 452 LSEPTMOPICVGLVYLAQMGSG-----KSFLTKLIEDEMT 486
Db 995 LTSPLE--LAAGSANNAGDGAAGTERKEGEVAGLDNNISQFLSLSGIETHRIDFET 1049

```

RESULT 5

Sequence 2 Application US/09841835
Patent No. US20020076795A1
GENERAL INFORMATION:
APPLICANT: de Lange, Titia
APPLICANT: Smith, Susan
TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRP1 AND METHODS
TITLE OF INVENTION: OF USE THEREOF
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue, 4th Floor
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/841,835
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/196,387
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-230 CIP1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1327 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-09-841-835-2

Query Match	4.58;	Score 395.5;	DB 10;	Length 1327;
Best Local Similarity	23.38;	Pred. No. 1.8e-16;		
Matches 167;	Conservative 80;	Mismatches 211;	Indels 259;	Gaps 17

QY 1 EEEINIPALKALLCEKCDYVERNECQOTPLMLAEGANEIYEKLLKNCANLEBDWMT 73
 348 EKKIMALLPLWNCNCHSDGRK ---STPLHLAAGRNVRITQYLLQGHADVHADNKGGLV 400
 QY 74 ALLSASKEHHIVVELLSGASLEHRDGMGTALMLACCYGRDYVELLSHGANSV- 133
 405 PLHMASCYGNYUTELLKNGACVNAAMDQCTPLPHEASNNRVEVCSLLSHGADPLV 466

QY 133 -----T 133
Db 465 NCHGKSAVDMAPFPELRELTFFKSHLLQARADLAKVKTALLETINKQPSHET 524
QY 134 GLOYSV-----YPIIWAAGRHADIVHLLONGARV 164
Db 525 ALHCVAASLHPRKQYOTELLFRKANVNEKNDFMTPLHVAABRAHNDVMEVLFHKGAKM 584
QY 165 NCSDKGTTPLVAAKRGHECYKHLAMGADVDQSGANSMTA-----207
Db 585 NALDTLGQTALHRAALAGHLQTRLLSTGSPTSISLOGFTAAQNGNEAVQOILSESTP 644
QY 208 -----LIVAKGGYTQSVKEILKRNPNVNLTKDG--NTALMIASKEGHEIEYODL 256
Db 645 IRSDVDYRLLESKAGDLETKOLCS--SONVNCRLBEGHSHPLHFAAGYNVSYVEL 703
QY 257 LDAGTYVNPDRSGDYLIGAVRGHVEIVRALLQKYADIDIRGQDNKTALYAVKGG-- 314
Db 704 LHHGADVHAHDKGGLVPLHNAQSYGHVEVALLVRGASAVNADLWKFPLHBAAKGKY 763
QY 315 -----NATWVRILQ-----C 325
Db 764 EICKLLIKGADPTKKNRDNPTLDLVKESGDTIDILKGDALDLDAKKGLARVQKLC 823
QY 326 NPDEICTKD---GEPPLIKATKRNIEVELLEDKAKYSAVDKRPPLH-----374
Db 824 TPEINCRDTQGRNSPTPLHAAGYNNLEVAEYILLEHGAADVNAQDKGLPLHNAASYGHV 883
QY 375 -----VAIRGSRRLAEILLRNPKGRLLYRPNKAGETP 408
Db 884 DIALALKIKYCNVATDKMAFTPLHBAOKGR--QICALLIAGADPTM---KNQSGQTP 939
QY 409 YN-----IDCSHOKSILT-----QIFGARHLSPTEPDGMDGLYDLKSSALADI 451
Db 940 LDATADDDIALLIDAMPPEALPTCKEQATVVASLISAPSPS-----CLSAASSIDN 994
QY 452 LSEPTMOPICVGLYAQWQSG-----KSFLLKLEDEMKT 486
Db 995 LTGRLAE--LAVGASASNAQGAAGTERKEGEVAGLDMNISQFLKSLGLEHLRIPEFT 1049

RESULT 10
US-09-835-788A-18
; Sequence 18, Application US/09835788A
; Patent No. US2002007458A1
; GENERAL INFORMATION:
; APPLICANT: NI et al.
; TITLE OF INVENTION: Death Domain-Containing Receptor Polynucleotides, Polypeptides, and
; FILE REFERENCE: PTO18P1
; CURRENT APPLICATION NUMBER: US/09/835, 788A
; CURRENT FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: PCT/US00/28666
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/159,585
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 60/167,246
; PRIOR FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 18
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-835-788A-18

Query Match 4.4%; Score 389.5; DB 10; Length 285;
Best Local Similarity 37.8%; Pred. No. 4,2e-17;
Matches 101; Conservative 49; Mismatches 110; Indels 7; Gaps 5;

QY 43 MAAEGNVEIVKELKANGANCNLEDD--NMTALISKSGCHIHIVELLKSGASLEHRD 101
Db 1 MEASAGHVEAVRLDHHGAGINTHSNEFESALLTACVKGHLDMVRLLEAGADQDHKT 60

QY 102 MCGWTALMACYKGRDVEVLLSHGANPSYTGLOYSVYPIIWAAGRHADIVHLLONG 161
Db 61 DEMHTALMEACMDGHVEAVARLLILDSGAQYNNPABDSFE--SPITLAAAGGHVLAALLIERG 119
QY 162 AKVNCSDRYGTPPLVMAARKHLECYKHLAMGADV--QEGANSMTALIVAAGGYQSV 220
Db 120 ANLEEVNDEGYPLMEAREGHEEMVALLIAGANINNOTBQETALTLLACGGSFEVA 179
QY 221 KEILKRNPNVNLTKDGNTALMIASKEGHEIEYODLDAQTYVNPDRSGDYLIGAVRG 280
Db 180 DFLIKAGADIEL---GCTPTLMEASQEGHLELVKYLKSLASGANVATTAATGDTALTACEN 236
QY 281 GHEIVRALLQKYADIDIRGQDNKTAL 307
Db 237 GHTVDADVLLQNGADLD--KQEDMKITL 262

RESULT 11
US-09-841-835-10
; Sequence 10, Application US/09841835
; Patent No. US20020076795A1
; GENERAL INFORMATION:
; APPLICANT: de Lange, Tilia
; APPLICANT: Smith, Susan
; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue, 4th floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/841,835
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/196,387
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-230 CIP1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 949 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-841-835-10

Query Match 4.4%; Score 387.5; DB 10; Length 949;
Best Local Similarity 25.2%; Pred. No. 3.6e-16;
Matches 144; Conservative 59; Mismatches 177; Indels 191; Gaps 11;

QY 14 EENETPALKALECKDVERNECCGOTPLMLAEOGNVEIVKELKANGANCNLEDDNMT 73
Db 348 EEKLMALPLPANNCHASDGR---STPLHLAGYVRVIVYDULLQHGADVHAHDKGGLV 404
QY 74 ALISAKRGHIEIVKELKANGASLEHRDMGWTALMACYKGRDVEVLLSHGANPSV- 132
Db 1 ALISAKRGHIEIVKELKANGASLEHRDMGWTALMACYKGRDVEVLLSHGANPSV- 132

Db 405 PLHNACSYGHEVTELLKHGACVANMDLMQFTPLHEAASKNRVEVCSLLSHGADPTLV 464
 QY 133 -----T 133
 Db 465 NCHGSAVDMADPTPELRERLYEFGKSHLQAAAREADLAKYKTLAETINFKOPOSHET 524
 QY 134 GLOYSV-----YPIIWAAGRHADIVHLLONGAKV 164
 Db 525 ALHCAVASLHPKRQVTELLKHGACVANNEKNKDFPTPLHVAERHNDVMEVLEKHGAKM 584
 QY 165 NCSDKYGTPLVMAARKGHECVKHLAMGADVDOGANSMTA-----207
 Db 585 NALDPLGOTLHRAACGHLQTCRLLLSYGSDPSTISLOGFTAAQMGNEAVOQLSESTP 644
 QY 208 -----LIVAAGGYTQSYVELLKNPNVNLTDKOG--NTALMSKEGHEIYVDL 256
 Db 645 IRTSDVDRLEASAGDELEVQKCS--SQYNCDLEGRSTPLHFAAGYNRVSVEYL 703
 QY 257 IDAGTYVNIPIRSGDTVLIGAVRGHEIVRALQKADIDIRGDNKTALYMAVEK-- 314
 Db 704 LHHGADVHAHAKRGGLVPLHNACSYGHEVALLVHGASVANVADLMKFTPLHEAASKKY 763
 QY 315 -----NATVRDIIQ-----C 325
 Db 764 EICKLLKHGADPTKKNRNDPTLDVKEGDTIDDLKGDPAALLDAAKGCLARVOKLC 823
 QY 326 NPDEICTKD--GETPIKATKMRNIEVEVLLDKGAKYSAVDKGDPTPLHVAIRGRSR 382
 Db 824 TPENINCRDPTGGRNSTPLHLAAGYNNLEVAEYLLHGHADVANAQDKGLIPLHNA-----877
 QY 383 RLAEILLRNPDKDRLLYRPNKAGETPNYIDC 413
 Db 878 -----ASYGCLARVOKLC--TPENINC 898

RESULT 12
 US-09-972-115A-6
 ; Sequence 6, Application US/09972115A
 ; Publication No. US20030032769A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Genon Corporation
 ; APPLICANT: Gregor, Morin B.
 ; APPLICANT: Walter, Funk D.
 ; APPLICANT: Mieczyslaw, Piatyszek A.
 ; TITLE OF INVENTION: A Second Mammalian Telomerase
 ; FILE REFERENCE: 080/003C
 ; CURRENT APPLICATION NUMBER: US/09/972,115A
 ; PRIOR FILING DATE: 2001-10-05
 ; PRIOR APPLICATION NUMBER: US 60/128,577
 ; PRIOR FILING DATE: 2000-04-10
 ; PRIOR APPLICATION NUMBER: US 60/129,123
 ; PRIOR FILING DATE: 1999-04-13
 ; NUMBER OF SEQ ID NOS: 64
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 6
 ; LENGTH: 1166
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-972-115A-6

Query Match 4.3%; Score 380.5; DB 9; Length 1166;
 Best Local Similarity 24.0%; Pred. No. 1,4e-15;
 Matches 157; Conservative 73; Mismatches 206; Indels 217; Gaps 13;
 QY 14 EENIPALKALLECKDVERNECGOTPLMLAEOGNEVIEVELLKGANCNLEDLDMWT 73
 Db 190 EKKMALLTPLNVNCHASDGRK---STPLHLAAGYNRKYVOLLQHGADVHAHAKGDDLV 246
 QY 74 ALISASKEGHIHIVEELLSGASLEHRDNGWTALMMAKYKRTDYVELLSHGANSPTV 133
 Db 247 PLHNACSYGHEVTELLVKGACVANMDLMQFTPLHEAASKNRVEVCSLLSYGADPTLV 306
 QY 134 G-----LQYS-----138

Db 307 NCHKSAIDLAPTPOLKERLAEFGKSHLQAAAREADYTRIKKHSLEWNEKHPQTHET 366
 QY 139 -----YPIIWAAGRHADIVHLLONGAKV 164
 Db 367 ALHCAASPYPRKQICELLKHGACVANNEKTEFELPLHVASBKANDVVEVVKHEAVY 426
 QY 165 NCSDKYGTPLVMAARKGHECVKHLAMGADV-----QEG-- 201
 Db 427 NALDNLGOTSLHRAACGHLQTCRLLLSYGSDPSTISLOGFTALQMGNEVQOLLQEGIS 486
 QY 202 -----ANSMTALIVAAGGYTQSYKE-----TL 224
 Db 487 LGNSEADROLLEAKAGDVEYVTKCTVGSVNCRDIEGRSTPLHFAAGYNRVSVEYL 546
 QY 225 KRNPNVNLTKDGNLTALMSKEGHEIYVDLLDAGTYVNIPIRSGDTVLIGAVRGHE 284
 Db 547 QHGADVHAHAKRGGLVPLHNACSYGHEVALLVHGASVANVADLMKFTPLHEAASKKY 606
 QY 285 IYRALLQKADIDIRGDNKTALYMAVEKGNATVRDIIQ-----324
 Db 607 ICKLLQHGADPTKKNRNDGTPV--DLVKGDTIDDLKGDPAALLDAAKGCLARVOKLC 664
 QY 325 CNPDEICTKDE--TPIKATKMRNIEVEVLLDKGAKYSAVDKGDPTPLH-----374
 Db 665 SSPDNVNC--RDPTGGRSTPLHLAAGYNNLEVAEYLLQHGADVANAQDKGLIPLHNAASYG 723
 QY 375 -----VAIRGRSRLAEILLRNPDKDRLLYRPNKAGE 406
 Db 724 HVDVAALLIKYNACVANATDKMAFTPLHEAASKRT--QLCALHLHGADPTLV--KNOEQO 779
 QY 407 TPNIDCSHOKSITLQFGARHLSPTEODMGYULYSALADIISEPTMP 459
 Db 780 TPDDIVSADVSAALLTRAMPSPALPSCYRQVILNGVRSFGATRDALSSPSSP 832

RESULT 13
 US-10-163-587A-15
 ; Sequence 15, Application US/10163587A
 ; Publication No. US20030096263A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Oliveira, Marcos
 ; TITLE OF INVENTION: SELECTIVE PARP-1 TARGETING FOR DESIGNING CHEMO/RADIO SENSITIZI
 ; FILE REFERENCE: 50229-306
 ; CURRENT APPLICATION NUMBER: US/10/163,587A
 ; PRIOR FILING DATE: 2003-01-10
 ; PRIOR APPLICATION NUMBER: 60/236,110
 ; PRIOR FILING DATE: 2001-06-07
 ; NUMBER OF SEQ ID NOS: 40
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 15
 ; LENGTH: 1166
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-163-587A-15

Query Match 4.3%; Score 380.5; DB 9; Length 1166;
 Best Local Similarity 24.0%; Pred. No. 1,4e-15;
 Matches 157; Conservative 73; Mismatches 206; Indels 217; Gaps 13;
 QY 14 EENIPALKALLECKDVERNECGOTPLMLAEOGNEVIEVELLKGANCNLEDLDMWT 73
 Db 190 EKKMALLTPLNVNCHASDGRK---STPLHLAAGYNRKYVOLLQHGADVHAHAKGDDLV 246
 QY 74 ALISASKEGHIHIVEELLSGASLEHRDNGWTALMMAKYKRTDYVELLSHGANSPTV 133
 Db 247 PLHNACSYGHEVTELLVKGACVANMDLMQFTPLHEAASKNRVEVCSLLSYGADPTLV 306
 QY 134 G-----LQYS-----138
 Db 307 NCHKSAIDLAPTPOLKERLAEFGKSHLQAAAREADYTRIKKHSLEWNEKHPQTHET 366
 QY 139 -----YPIIWAAGRHADIVHLLONGAKV 164

Db 367 ALHCAASPYPKRKOICELLIRKGANINKEKTEFLPLHVAASEKANDVEYVYKBAKV 426
QY 165 NCSDKYGTPLVMAARKGHLECYKHLANGADV-----OEG-- 201
Db 427 NALDNLGOTSLHRAAYCGHLOTCTRLLSYGCDDPNIISLQGFALQMGNNVQOLLQEGIS 486
QY 202 ---ANSMFALIVAVGGYQSVKE-----IL 224
Db 487 LGNSEADROLLEAAKAGDVEYKLCYQSVNCRDIEGRQSTPLHFAAGYNEVYVYELL 546
QY 225 KRNPNVNLTDKDGNTALMIASKEGHEIYODLLDAGTYVNIIPDRSGDTVLIGAVRGHVE 284
Db 547 QHGADVHAKDKGGLVPLHNAACSYGHEVAELLYKGAIVNVADLMKFTPLHAAAKGYE 606
QY 285 IYRALLQKADIDIRGQDNKTALYMAVEKGNATVNRDILQ----- 324
Db 607 ICKLLQHGADPTKKNRDNTPPL--DLVKDGDIDILQDGLDALLDAAKKGLARVKKL 664
QY 325 CNPDTEICTKDE---TPLIKATKMRNIEVEVELLDGAKVASVDKKGDTPLH----- 374
Db 665 SSPDNVNC-RDTQGRHSTPLHLAGYNNLEVAEYLQHGADVNAADKGLIPLHNAASYG 723
QY 375 -----VAIRGRSRRLAELLRNPKDGLRYPRNKAGE 406
Db 724 HVDVALLIKYNACVNAFTDKMAFTPLHAAQKGR-QLCALLLHAGADPTL---KNQEGQ 779
QY 407 TPNYIDCSHOKSILTOIFGARHLSPTETDGMIGYDLYSSALADLSEPTMOP 459
Db 780 TPDLVSADVVSALLTPAAMPSPALPSCYKPOVLNGVRSPQATDALSSGSPSP 832

RESULT 14
US-09-509-196A-2
; Sequence 2, Application US/09509196A
; Patent No. US20020037582A1
; GENERAL INFORMATION:
; APPLICANT: DAIRY, Roger J.
; APPLICANT: SOTHELAND, Robert L.
; TITLE OF INVENTION: A Potential Effector for the Grb7 Family of Signalling
; FILE REFERENCE: 1871-129
; CURRENT APPLICATION NUMBER: US/09/509,196A
; CURRENT FILING DATE: 2000-03-23
; PRIOR APPLICATION NUMBER: P09388
; PRIOR FILING DATE: 1997-09-23
; PRIOR APPLICATION NUMBER: PCT AU98/00795
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1074
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-509-196A-2

Query Match 4.2%; Score 371.5; DB 10; Length 1074;
Best Local Similarity 23.7%; Pred. No. 4.5e-15;
Matches 155; Conservative 73; Mismatches 208; Indels 217; Gaps 13;
QY 14 EENENIRALKALLEKCDVBERNCCGOTPLMLAEGGNVEIVKELKNGANCNLEDIDNNT 73
Db 98 EENMMLLPPLVNCASDGRK---STPLHLAGYNNRIVYDILLQHGADVNAKDKGDLV 154
QY 74 ALISASKEGHIHIVEELKSGASLEHRDGMGTALMAGCYKGTVDVVELLSGANGSPVT 133
Db 155 PLHNAACSYGHEVTELLVHGGCVNADMLOFTPLHAAKSNRVEVCSLLSYGADPTLL 214
QY 134 G-----LQYS----- 138
Db 215 NCKNSAIDAPLPOLKERLAYEFKSHSLQARADVTYRIKKHLSLEWVNFKNPQTHET 274
QY 139 -----VPIIWAAGRHADIVHLLONGAKV 164

Db 275 ALHCAASPYPKRKOICELLIRKGANINKEKTEFLPLHVAASEKANDVEYVYKBAKV 334
QY 165 NCSDKYGTPLVMAARKGHLECYKHLANGADV-----OEG-- 201
Db 335 NALDNLGOTSLHRAAYCGHLOTCTRLLSYGCDDPNIISLQGFALQMGNNVQOLLQEGIS 394
QY 202 ---ANSMFALIVAVGGYQSVKE-----IL 224
Db 395 LGNSEADROLLEAAKAGDVEYKLCYQSVNCRDIEGRQSTPLHFAAGYNEVYVYELL 454
QY 225 KRNPNVNLTDKDGNTALMIASKEGHEIYODLLDAGTYVNIIPDRSGDTVLIGAVRGHVE 284
Db 455 QHGADVHAKDKGGLVPLHNAACSYGHEVAELLYKGAIVNVADLMKFTPLHAAAKGYE 514
QY 285 IYRALLQKADIDIRGQDNKTALYMAVEKGNATVNRDILQ----- 324
Db 515 ICKLLQHGADPTKKNRDNTPPL--DLVKDGDIDILQDGLDALLDAAKKGLARVKKL 572
QY 325 CNPDTEICTKDE---TPLIKATKMRNIEVEVELLDGAKVASVDKKGDTPLH----- 374
Db 573 SSPDNVNC-RDTQGRHSTPLHLAGYNNLEVAEYLQHGADVNAADKGLIPLHNAASYG 631
QY 375 -----VAIRGRSRRLAELLRNPKDGLRYPRNKAGE 406
Db 632 HVDVALLIKYNASINATDKMAFTPLHAAQKGR-QLCALLLHAGADPTL---KNQEGQ 687
QY 407 TPNYIDCSHOKSILTOIFGARHLSPTETDGMIGYDLYSSALADLSEPTMOP 459
Db 688 TPDLVSADVVSALLTPAAMPSPALPSCYKPOVLNGVRSPQATDALSSGSPSP 740

RESULT 15
US-09-972-115A-2
; Sequence 2, Application US/09972115A
; Publication No. US20030032769A1
; GENERAL INFORMATION:
; APPLICANT: Geron Corporation
; APPLICANT: Gregg, Morin B.
; APPLICANT: Walter, Funk D.
; APPLICANT: Mieczyslaw, Piatyszek A.
; TITLE OF INVENTION: A Second Mammalian Telomerase
; FILE REFERENCE: 080/003C
; CURRENT APPLICATION NUMBER: US/09/972,115A
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: US 60/128,577
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: US 60/129,123
; PRIOR FILING DATE: 1999-04-13
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1333
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: The 'xaa' at location 1 stands for Lys, Asn, Arg, Ser, Thr, Il
; OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon
; OTHER INFORMATION: yr, Ttp, Cys, or phe.
; NAME/KEY: misc.feature
; LOCATION: (2)..(2)
; OTHER INFORMATION: The 'xaa' at location 2 stands for Lys, Asn, Arg, Ser, Thr, Il
; OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon
; OTHER INFORMATION: yr, Ttp, Cys, or phe.
; NAME/KEY: misc.feature
; LOCATION: (3)..(3)
; OTHER INFORMATION: The 'xaa' at location 3 stands for Lys, Asn, Arg, Ser, Thr, Il
; OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon
; OTHER INFORMATION: yr, Ttp, Cys, or phe.
; NAME/KEY: misc.feature
; LOCATION: (4)..(4)


```
OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Glu, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (34)..(34)
; OTHER INFORMATION: The 'xaa' at location 34 stands for Lys, Asn, Arg, Ser, Thr, Ile,
; OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Glu, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (35)..(35)
; OTHER INFORMATION: The 'xaa' at location 35 stands for Lys, Asn, Arg, Ser, Thr, Ile,
; OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Glu, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (36)..(36)
; OTHER INFORMATION: The 'xaa' at location 36 stands for Lys, Asn, Arg, Ser, Thr, Ile,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.

Query Match      4.1%; Score 360.5; DB 9; Length 1333;
Best Local Similarity 26.5%; Pred. No. 3.2e-14;
Matches 137; Conservative 69; Mismatches 198; Indels 113; Gaps 12;

OY 14 EENIPALKALKCKDVERNECGOTPLMLAEGCNVEIVKELKNGANCNLEDLDNNT 73
Db 348 EKKMMLTPLNCHASDRK--STPLHLAGYNKIVQLLQHGADVHAKDKDLY 404
OY 74 ALISASKEGHIIVEELKSGASLEHRDMGWTALMACYKGRDVEVLLSHGANPSY- 132
Db 405 PLHMACSYGHEYTELVKAGACYNAMDLMQFTPLHBAASKNRYEVCSLLSYGADPTLL 464
OY 133 -----TGLQSVYPIIWAARG-----HAD 152
Db 465 NCHNKAIDLAPTPOLKERLAYEFKGSLLQAAAREADVTRIKKHSLEWVNFKHPOTHE 524
OY 153 IVH-----LLONGAKVNCSDKGTTPVMAARKGHLECYKHLAMGADY 197
Db 525 ALHCAASPYKPKKQICELELLRKGANINEKTEFELTPLHVASEKANDVEVVKHEAKY 584
OY 198 DOEGANMTALIVAVKGYTSQVKEILKRNPNVNLTDKGNFTALMTASKGHEIVODLL 257
Db 585 NALDNLGQTLHRAAYGHLQTCRLLSYGCDPNIISLQSGFTALOMGN-----ENVQQL 639
OY 258 DAGTYVNIPIRSGDYVLIGAVRGHVEIVRAL--LOKYADIDIRGODNKATLYAVEKGN 315
Db 640 QEG--ISLGNSEADRLLEAKAGADVETVKLCTVQSVNCRDIEGRQS--TPLHFAAGYNR 696
OY 316 ATWVRDLQCNPTETICTKGETPLIKATKRNIEVEVLLLDGAKYSAVDKKDTPPLHY 375
Db 697 VSVVEYLLQAGADVHAKDKGLVPLHMACYXGHEVAELLVKHGAVVNVADLMKFTPLHE 756
OY 376 AIRGRSRRRLAELL-----RNPGRGLLYRPKAGET-----407
Db 757 AAKKGYEICKLLQHGADPTKKNRGNNTLLDLVKGXCTIDIOXLRGDVALLDAKKGCL 816
OY 408 -----PYNIDCSHQSILJOLFGARHLSPETEDG 436
Db 817 ARVKKXXFPDNNCRD-----TQ---GRHSTPLHLAG 845
```

Search completed: July 1, 2003, 14:50:57
Job time : 45 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2003, 14:34:05 ; Search time 19.5 Seconds
(without alignments)
2587.708 Million cell updates/sec

Title: US-10-021-571-2

Perfect score: 8853

Sequence: 1 MSVLISQSVINVEENIPAAAA.....ELHAASSTGTGFERESTL 1715

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Database: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTCUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/PTCUS.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	536	6.1	1839	2	US-09-172-977-4
2	503.5	5.7	843	2	US-09-172-977-3
3	466.5	5.3	1745	2	US-09-031-485-33
4	466.5	5.3	1745	2	US-08-847-429A-33
5	466.5	5.3	1745	3	US-09-065-474-33
6	466.5	5.3	1745	3	US-09-065-474-33
7	449.5	5.1	1088	4	US-09-082-059-2
8	414.5	4.7	1423	4	US-08-810-712-10
9	395.5	4.4	1327	4	US-09-196-387-2
10	387.5	4.4	949	4	US-09-031-485-28
11	376	4.2	348	2	US-08-847-429A-28
12	376	4.2	348	3	US-09-065-474-28
13	376	4.2	348	3	US-09-065-474-28
14	376	4.2	348	4	US-09-557-034-28
15	373	4.2	1166	4	US-09-350-982C-5
16	359.5	4.1	352	3	US-09-065-474-139
17	359.5	4.1	352	3	US-09-557-034-139
18	357.5	4.0	741	2	US-08-436-771-4
19	357.5	4.0	741	2	US-08-436-998-4
20	357.5	4.0	741	2	US-08-436-998-4
21	357.5	4.0	741	2	US-08-487-797-4
22	357.5	4.0	741	2	US-08-487-797-4
23	356	4.0	679	2	US-08-436-771-6
24	356	4.0	679	2	US-08-436-998-6
25	356	4.0	679	2	US-08-487-797-6
26	356	4.0	679	2	US-08-701-005A-4
27	356	4.0	679	2	US-08-479-895-4

28	356	4.0	679	5	PCT-US95-02058-6	Sequence 6, App1
29	355	4.0	787	4	US-09-188-930-334	Sequence 38, App
30	353	4.0	302	2	US-09-031-485-38	Sequence 33, App1
31	353	4.0	302	2	US-08-847-429A-38	Sequence 38, App1
32	353	4.0	302	3	US-09-065-474-38	Sequence 38, App1
33	353	4.0	302	4	US-09-557-034-38	Sequence 38, App1
34	353	4.0	303	2	US-09-031-485-23	Sequence 23, App1
35	353	4.0	303	2	US-08-847-429A-23	Sequence 23, App1
36	353	4.0	303	3	US-09-065-474-23	Sequence 23, App1
37	353	4.0	303	4	US-09-557-034-23	Sequence 23, App1
38	352	4.0	679	3	US-08-943-956A-4	Sequence 4, App1
39	347.5	3.9	741	2	US-08-462-481-2	Sequence 2, App1
40	347.5	3.9	741	2	US-08-436-771-2	Sequence 2, App1
41	347.5	3.9	741	2	US-08-436-998-2	Sequence 2, App1
42	347.5	3.9	741	2	US-08-487-797-2	Sequence 2, App1
43	347.5	3.9	741	2	US-08-701-005A-2	Sequence 2, App1
44	347.5	3.9	741	5	PCT-US95-02058-2	Sequence 2, App1
45	347.5	3.9	741	5	PCT-US95-02058-2	Sequence 2, App1

ALIGNMENTS

RESULT 1
US-09-172-977-4
; Sequence 4, Application US/09172977
; Patent No. 5989663
; GENERAL INFORMATION:
; APPLICANT: Tany, Y. Tom
; APPLICANT: Guebler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Yue, Henry
; TITLE OF INVENTION: HUMAN ANKYRIN FAMILY PROTEIN
; FILE REFERENCE: PF-0615 US
; CURRENT APPLICATION NUMBER: US/09/172,977
; CURRENT FILING DATE: 1998-10-14
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PERL Program
; SEQ ID NO 4
; LENGTH: 1839
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: g29491
US-09-172-977-4

Query Match	6.1%	Score 536;	DB 2;	Length 1839;
Best Local Similarity	19.7%	Pred. No. 1.4e-35;		
Matches 398;	Conservative 276;	Mismatches 706;	Indels 636;	Gaps 70;
14	EEENIPALKALEKCKDYDERNEGQPTPLMAEOGANEIYKELLKNGANCLEDDNMT	73		
73	KSGHGVYQELLGRSSVDSATKKNNTALHIASLAGAEVYKLVKEGANINMAOSQNGFT	132		
74	ALISAKGKHIVYELLKSGAS-----LEHRDMG---	103		
133	PLYMAOEHHIDVYKLLNGANOSTATEDGFTPLVALOOGHNOVAALLLENPTKGVKVR	192		
104	-----GWTALMAACVGRDGVYELL	123		
193	LPALHIAKDKDTKSAALLLONHDNADYOSKMMVNRTEGTPPHIAHAGNVAVTL	252		
124	LSHGAPSVYGLQYVYPIIMAGRGHADIYELLONGAKYVCSKYGTPLVMAARKG-	182		
253	LMRGAVPETA-RNGITPLHVASKSGNTNMVYLLDRGGQIDAKTRDGLTLHCAARSG	311		
183	-----HLECYHLLAMGADVDOEGANSKTLALIV	210		
312	DQVVELLEERGAFLARTKNGLSPLHMAAGDHVCEVHLLHOKAPVDVTLDYTLALHV	371		
211	AKVGCTGVSKRLKRNPNVNLTDKGNATLMAISKE-----	247		
372	AAHGHYRVTKLLDRANPNARALNGFTPLHIAKKNRIKVMELLYKGIASIQAITESG	431		

QY 248 -----GHIEYODLLDAGTYVNIIPDRSGDTVLIGARGGHEIVETRALLOKXADID 297
 Db 432 LPIHVAAPMGHNLIVLLLONGASPDVTNIRGETALMAARAGVEVAVRCLRGALVD 491
 QY 298 IRGONKTAALYWAVERGNATVNDILOCPTEICTKDETPLIKATMRNIEVELLD 357
 Db 492 AARAEQTPHIASTRGKTEIYQOLLQMAHPDATTNGYTPHISAEGGVDAVYLE 551
 QY 358 KGAKSAVDKGDTPHVAIRGRRLABELL-----NEKD 394
 Db 552 AGAHSATKKGEPHVAARYGSLDVAKLLQRRAAASAGKNGLPPLHVAHYDMQV 611
 QY 395 GRLL-----YRPKAGETPYNIDC-SHOKSILTOI--FGAR-----HLSP 432
 Db 612 ALLLEKASPHATKNGYTPHIAKKNQMOIASILLNTYGAETNITYKQGVTPHIAAQ 671
 QY 433 ETDGML-----GYDLYSALADILSEPTMOPICVGLVQMGSGSKFLKLEDEM-- 484
 Db 672 EGHDMVTLDDKGNHIMSTRSGITS-----LHLAAQ-----EDKYNV 710
 QY 485 -----KTFAGOOTEPFLQFQSMILVFLLLC--GGILGY-PAFVDTMLAIAISLPLAL 536
 Db 711 ADILTRGADODAHKLGITPLIV-----ACHYGVKKNVFLKQAGAVNAKTKNGYTP 765
 QY 537 -----TYFFIYIFGRRGESENMAMALSTRLARHIGYLELLFKLMEVNP----- 584
 Db 766 HOAAQOGHTHIIIVLLOHGAKPNATTANGNTALA--IAKRGYISVDTLKVIVETTT 823
 QY 585 -----ELPRTTALPYRPLFYDYNLSVGGSTLAEMITAILSDACEREGPL- 633
 Db 824 TTTTEKHKNVPELMEIVLDVSDDEGD-DTMTGDGEYLPEDLKEIGDLSPPSOFID 882
 QY 634 -----ATRLFRVPR-----EE 645
 Db 883 GMYNIRYSLGEGRSLSRFSSSDHSILSHASTYLDASVMDSVYIISHOVSTLAKAER 942
 QY 646 SGGKKW-----KTYCCLPSFVIFLTV-----GCIAGITLLAIFRVDPRKL 688
 Db 943 NSYRLSMGTENLDVVALSSPIHSGFLVIFWVDARGAMRCGRHNGRLTI--IPPRKC 998
 QY 689 TVNAILISASVGLAFVPLNRTMVOYDLSLNSQKRL-----HSAKSLHLKLS 739
 Db 999 TAPF-----RYTCRL-----VKRHLATMPVGEGLASRLIEVGP 1035
 QY 740 EG-----FMVLKCEVELMARMAKTIIDSFON-----OTRLVYIIDGLACE 781
 Db 1036 SGAGFLGPVIEIHFALRKREBELVYLRSENDSKHEKCDYTEDELIELNGMD--- 1092
 QY 782 QDKVLOMDIVRVLESGPFIATASDPH--IIRKAINOLNSVLRSDNINHDMYRNI 839
 Db 1093 -----EVLDSPEDELEKRIKRIIRDEPQYFAVYSRIQ-----DSNLIGPE----- 1134
 QY 840 HLPFLNLSRGLSNARKFLVTSAT--NGDICSDDTGOEDTDRVSONSLGEMTK--LGS 895
 Db 1135 -----GVLSTSTVYPOVQAVPEBALTKRIRIVGQ-----AOPMHESELVKILLGN 1179
 QY 896 KTALENRDYV--RRROMQRTTROMSEFLTKLVTEDMFSDISFOTMRRLNIVSVGR 953
 Db 1180 KATESPIVLEPRRRKFKHPTMTWIPVKAASDVMLNGFGDAP-TIARLL--CSITGCT 1235
 QY 954 LRAQITFNDRLASWINLT--EOMPYRTS--WLIIYEETESLPDMOTLKTYER 1005
 Db 1236 TPA-----QWEDITGTPLEFVNCEVSTTNVSAFWLI-----DCRQIQESVTFASQVYR 1286
 QY 1006 ISKNIPTTKOYEPLELIDGDIRNEFVLSRTPIVARADVTKPLPCVINLDPKL----- 1059
 Db 1287 EITCVPR-----MAKVFAKSHD-----IARIRCGMCTDDVDVKTLEQ 1327
 QY 1060 REIADYRAAREQINIGLAVPPLPLHECPRPSPSGYQAPASVCSASFNGPPGVVSP 1119
 Db 1328 QENAEVARSMD-----VEVLEGRPIYVDCGNLVPLTKSG----- 1363

QY 1120 QPHSSYSGLSGPHFPYNAAPATGSSILLSSM--TVDVCEKL-----ROISETDON 1172
 Db 1364 -----QHIFSFPAKENRLPLEVAVRDTIOEPCGRSLSEKPEKSTRGVHQ 1410
 QY 1173 MPOYCTTIKANINGRVLSOCNIDELKEMANFDMHILFSMVLSEKRSVOVPEDP 1232
 Db 1411 -----ALONMILITPIYTKESSEQOE-----EELDMTSSEKPODE 1447
 QY 1233 RFLNENSAVPH-----GESARRSHHELPLELISOTPYTLNFSPELNLGLDEGA 1286
 Db 1448 QERIEERLAIADHLGFSWTELARELDFEEQIHQIRINPNLSLODOSYLKILWELRG 1507
 QY 1287 PRHSNLSWOSQTRPTSLSSNGDSSTESKTLTDVQAEYDAVEYTAQMSOLEGCTG 1346
 Db 1508 -KIA-----TDVNLVECLTRINMDVHLMETNTEPLQERISHSYAE-TEOTITLDSHG 1560
 QY 1347 SSTISGR-----SPHSYTYIGOSSGGSINHT 1374
 Db 1561 FSVLOEELCTAOKQKEQAVSKESPTCDHPPIVSEEDISVGFQDGVPRTEGDSST 1620
 QY 1375 LEQERKEGELKQEDGRKSTLMKRGVIDYISSGVSTNAPSLDPTTEDEKSDSGSKL 1434
 Db 1621 ALFPQTHKEQVQODFSGKMODLPRESSLEYQOEYVT-----TPGETSETOKAMI 1671
 QY 1435 LPG--KKSSEPSLFTQDLKLGGLRYOKLPDSDESGTGRVQITPHCSKMIIRRLKA 1492
 Db 1672 VPSSPKTPEVSTPAEERKL-----YLOTPSSRSGSPILO----- 1709
 QY 1493 KORECASPOHSAE--PIRTFIKAEYUSDALLDKKSDSGVRSNPNHSLNEAD 1550
 Db 1710 ---EPEPSHRESSPRKTSIYIV-----SADNOPENCERLDEDAFAEKGD 1755
 QY 1551 DSQLEKANLIE---LEDBGHSGRGMPSHLSGLDPTIARMSCEDKSPSECLIAS 1607
 Db 1756 MPETPEYITEEYIDEGHTYVKKVTKR-----IIRRVSSBG--TEKEELIMVQOM 1805
 QY 1608 PEESSMPA--COKAVNLNRPSTYTLNNTAPTNRANO 1642
 Db 1806 PQE--PVNIEEGDYSKVIKRVYLSKDTQSDENNE 1839

 RESULT 2
 US-09-172-977-3
 ; Sequence 3, Application us/09172977
 ; Patent No. 5989863
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang, Y. TOM
 ; APPLICANT: Guegler, Karl J.
 ; APPLICANT: Corley, Neil C.
 ; APPLICANT: Yue, Henry
 ; TITLE OF INVENTION: HUMAN ANKYRIN FAMILY PROTEIN
 ; FILE REFERENCE: PF-0615 US
 ; CURRENT APPLICATION NUMBER: US/09/172,977
 ; CURRENT FILING DATE: 1998-10-14
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 3
 ; LENGTH: 843
 ; TYPE: PRT
 ; ORGANISM: Rattus norvegicus
 ; FEATURE:
 ; OTHER INFORMATION: g1841966
 US-09-172-977-3

 Query Match 5.7%; Score 503.5; DB 2; Length 843;
 Best Local Similarity 24.6%; Pred. No. 1,8e-33;
 Matches 203; Conservative 110; Mismatches 273; Indels 239; Gaps 20;

 QY 14 EENIPALKALKKCKVDNERNECGOTPLMLAEOGNEIYKELLKNGANCNLELDQMT 73
 Db 38 KEGHGVIGVQELLGRSSVDSATKKNRTALHLIASLAGAEEYKVLKBEKANINMAOSQNGFT 97
 QY 74 ALISASKEGHIHIVEELKSGAS-----LEHRDMG--- 103

```

Db 98 PLYMAQENHIDVYKLLNGANQSTATEDGFTPLAVALQGHNOAVALLNDKRGYR 157
QY 104 -----GFTALMACKRTDVEILL 123
Db 158 LPAHIAARKDDTKSALLQNDHNDVOSKMMVNTTESGFTPLHIAHNGVNAVATLL 217
QY 124 LSHGANPSVTGLQSVYPIIMAAGRGHADIYHLLQNGKAVCSQYGTTPYVMAARG- 182
Db 218 LNRGAADVFTA-RNGITPLHVASRKGNTMMVKLLDRGQIDAKTRDGLTPLHCAARGH 276
QY 183 -----HLECVKHLAMGADVDOEGANSTALIV 210
Db 277 DQVVELLERGAPLARTKNGSLPILMAAGSHVEYKHLQHKAPVDVITDYLTAHV 336
QY 211 AVKGYTSVKELKRNPNVNTDKDGNALMASKE----- 247
Db 337 AAHGHRVYTKLLDRANPNARALNGFTPLHACKNRKRYKVELLYKGYAIOAITESG 396
QY 248 -----GHIEIYODLLDAGTYVNIIDRSQDYLIGAVGGHVEIYRALLQKAYDID 297
Db 397 LPTPYAATFMGHLNITVLLQNASPDVTNIRGETALHMAARGVEYVRCULRNGALVD 456
QY 298 IRGQDNKTLALYMAVEKGNATMYRDIQCNPDTEICTKDETPLIKATKRNIEVEILLD 357
Db 457 ARAREGOTPLHIASRIGTEIYVQLQGHMAHPDAATTNGTPLHISARGQYDVASVLE 516
QY 358 KGAQVANDKGDTPPLHVAIKGRSRLAELLR-----NPKD 394
Db 517 AGAHSLATKKGFTPLHVAAKYGSIDYAKLLQRRAAADSAGKNGLTPHVAHYDNOFY 576
QY 395 GRLL-----YRPNKAGETPNYNDI-SHOKSILTOI-FGAR-----HLSP 432
Db 577 ALLLEKGSAPATAKNGTPLHIAKKQMOJASTILNGAFTNTVTOGVTPLHLSQ 636
QY 433 ETDGMLGYDLYSSALADILSEPTMOPRICVGLYAGWGSQKSL-LKLEDE----- 484
Db 637 EGHDMVTYVLEKGA-----NIMSTKSGLTSLHIAEDKYNVADILT 680
QY 485 KTAGAGOTPLPFPMSIYVLTLLC--GGGLGV-FAPVDYMLAIAISLAL----- 536
Db 681 KHGADDDATYTKLGYTPLIV---ACHYGNVKKVNFLLKOGAVNNAKNGYPLHQAQ 735
QY 537 ---IYFPIYVFGREGESMMAMALSTRLARHIGYELLFKLMEVNP----- 584
Db 736 QGHRHIIIVLQGAKRNATTANGNTRALA--IAKRLGYSVDTPLKAVVEYTTTTTIT 793
QY 585 -----ELPEQTKALPYRFLFTDYNRLSSVGETSLAEMIATLSD 624
Db 794 EKHLNAPETMTETVLDVSDEGD-DVYTGDSGEYLRPELDKELGD 837

```

```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: WordPerfect for Windows, Version 7.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/031,485
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/847,429
ATTORNEY/AGENT INFORMATION:
NAME: Verser, Carol Talkington
REGISTRATION NUMBER: 37,459
REFERENCE/DOCKET NUMBER: HW-5
TELECOMMUNICATION INFORMATION:
TELEPHONE: 970/493-7272
TELEFAX: 970/484-9505
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 1745 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-031-485-33

Query Match
Best Local Similarity 20.4%; Score 466.5; DB 2; Length 1745;
Matches 358; Conservative 235; Mismatches 575; Indels 591; Gaps 73;

14 EENIPALKALLEKCKDYDERNEGOTPLMAEGGNEIYKELKNGANCLEDDN-- 71
207 KKDPTKATLTLQNHNSDVTSSGFTPLHIAHNGVNAVQALLLEKGANVYARHNIS 266
72 -----W-----TALISAKRGHIIYELLKSGASL-- 97
267 PLHVAATKGRIMVSLLAHGAVIDCRTLITPLHCASRSGHDOVDDLEKAPISAK 326
98 -----EHRMGWTLAMW 110
327 TNGIAPLMAQVDDVYDITPLHVAHGHVAVAKLLDRNADPNARALNGFTPLHI 386
111 ACYGRTDVVELLSHGANSVYGLQSVYPIIMAAGRGHADIYHLLQNGAKYNSDKY 170
387 ACKKNRIKIVLELLYHYAIEAT-TEGSLPLHVAAFGAINIYVLLQOGANADVAVR 445
171 GTPLVMAARKHLECVKHLAMGADVDOEGANSTALIVAKGYTSVKELKRNPNV 230
446 GETPLHARANOTIVRYLVNRNGAOVDAARELOTPLHISRLGENTDVIILLQANSP 505
231 NITDKGNATLMAASKEGHIEIVODLDAGTYVINIPDRSGDTVLIGAVRGHVEIVRALL 290
506 NATRDLYTPHIAKKEGEVYALILMDHGTCKTLTKKGFPLHIAKNGULPYAKSL 565
291 QRYADIDIRGQDNKTLALYMAVEKGNATMYRDIQCNPDTEICTKDETPLIKATKRNIE 350
566 ERGTPEVDIEGKNQVPLHVAHAYNNDKVALLLLENGASAHAAKNGYPLHIAKKQMD 625
351 VVELLDKAKVASVDKGDTPPLHVAIGRSRLAELLRN----- 391
626 IASTLTHYANANAESKAGFTPLHAAQGHHEMALILENGAKKGAOARGSLTPMHCLA 685
392 -----PRDGRILYRPNKAGETPNYNDISHOXSILTOIFGARHLSPTETDGMIG 440
686 QEDRVSAVELYKENNAIDPKTKAGTTPHVAACH-----FG-----QINMYR 727
441 YDIYSALADILSEPTMOPRICVGLYAGWGSQKSLKLK-----EDMKTPFAGOTPEPL 495
728 FLEHGARVSVITRASVTP-----LHOAAQGHNSVRYLIEHGASPNVHTSTGO----- 777
496 FQESWILVLTLLCGGLVFAFPVDYMLAIAISLALIIYFPIYVFGREGESW 555
778 -----TPLST----- 782

```

RESULT 3
 US-09-031-485-33
 Sequence 33, Application US/09031485
 Patent No. 5824306
 GENERAL INFORMATION:
 APPLICANT: Tang, Liang
 TITLE OF INVENTION: DIOFILARIA AND BRUGIA ANKYRIN
 TITLE OF INVENTION: PROTEIN, NUCLEIC ACID MOLECULES, AND
 NUMBER OF SEQUENCES: 85
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Carol Talkington Verser, Ph.D.
 ADDRESSEE: Heska Corporation
 STREET: 1825 Sharp Point Drive
 CITY: Fort Collins
 STATE: Colorado
 COUNTRY: USA
 ZIP: 80525
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk

Db 506 NAATRDLYPPLHIAKEGEVEAAIIMDGTDLTKKGPPLHIAAKYGNLPVAKSL 565
 QY 291 OKTADIDIRGODKKTALYMAVEKGNATMVDIIQCPDIEICIKDEPLIKTKRNI 350
 Db 566 ERGTPVIEGKNQVTPPLVAHAYHNKVALLLLENGASAAHAKNGYTPPLHIAKNOMD 625
 QY 351 VVELLDKGAASAVDKGDTPLVAIRGSRRLAELLRN----- 391
 Db 626 IASTLLHYKANANAESAGTTPPLHIAQEGHREMAALLIENGAKVGAQARNGILPMHLCA 665
 QY 392 -----PKDGLRYPNKAQETPNYIDCSHOKSILITQIFGARHLSPTETDGMIG 440
 Db 686 QEDRVSAEELVENAIDPKTAGYTPPLVACH-----FG-----QINMYR 727
 QY 441 YDLSSALADLSEPTMOPPICGIVAQWSSGSKFLKT-----EDEMKTAGOOTEPL 495
 Db 728 FLIEHGARSVITRASTP-----LHOAAOQHNSVRYLLEGASPDVHTSGO----- 777
 QY 496 FOFSWLIVFTLLCGIGLVFAPVPVNTLAISLFLALYIFIVIFGGRRGESW 555
 Db 778 -----TPLSI----- 782
 QY 556 NMAALSTRARHIGYELLFKIMFVNPPPLPEOTTKALPVR-----LFTDYN 604
 Db 783 -----AERLGYSVEALKITETVITETVETTERYKRPONPEAMNETMFS-- 830
 QY 605 RLSSVGETSLAEMIALDSCEREGFL--ATRLRFPVPTESGOKK----- 650
 Db 831 --SEDEGEDNOITANAHAFSESILKGLHDSIGVHILHATEPTLSKSPVEBTDGDLA 888
 QY 651 -----KMKTCCLSPFVILFT-----VCIAGIT 676
 Db 889 LIRKAQHEPITTAADPSLDASLPDNTIMRTMQSPFLSPFVADANGAMRCRHSVGR 948
 QY 677 LLAIFRVPKHLVYNAIILISAYVGLAFLVLCRTMWOVIDSLN-----SQRRLHSAAS 732
 Db 949 IIPPRKAPQTRV-----TCR--YLGDKLHAPPLSEGEAL--AS 986
 QY 733 KIKHLKSEG--FMVKLCEVELAMARKATIDSFQONTIRLVII--DG-----LDAC 780
 Db 987 RILEMAHGAFLGFLVLEVPHPHRA--L-----RGREIVILRSDGQHMKEHQLEAT 1038
 QY 781 EODKVLQMLD-----TVRLFSGKPFALIASDPHIIKAIIONLNS 822
 Db 1039 E-DAVQEVLMNSPDAEELSOLDLHISRIIRILNDEPM--YEA-----VTVRQEVHC 1090
 QY 823 VLKDSNINGHDMYKNIV-HLPVFLNSRGLSNARKFELVTSATNGDITCSDTTGTQEDTRR 881
 Db 1091 VGPEGGV-----ILSSVPHVOAIFPDGSLTKTIKVSVOA-----QP 1127
 QY 882 VSQNSLGEMLTKL-GSKTALNRDITY--RRROMQRTITROKSPDLTK--LLVTEDMWSDI 935
 Db 1128 VPQET--VTRLHGNRAVSPVIVPEPRRRKFRHPTLCPLPOSSKGMLOYSGOPGQ 1184
 QY 936 SPQTRRLINIV-----SVTGRLLRANQITFNMRLASWINLEQ-W----- 976
 Db 1185 EPTLRLILCKSGTSSPAQWEDITG-----TTQULTFGEDEVSPITTVSARFWIMDCQPRD 1240
 QY 977 -----PRTSMILLYLEET--EG-----LPDQMLTKM--YERISKNI 1010
 Db 1241 AARMAOEYVNEALVAPPMARFLI-FARRTPAEQQLRFLCMTDREKTELEKOFTE-I 1298
 QY 1011 PTTKDVPEPL-----LEIDDIRNFVFLSSRPVLVARDVK--FLPCVIVN--LDPKL 1059
 Db 1299 AKSKDVEVLSGRHOFLFESGNL-----LPTYSQQLSLYFLPEQENRLAEVWKI 1348
 QY 1060 R-----EIIADVRAARQINIGLAVPPLDHEGPPRPSGQSPASVCSA--SFNG 1110
 Db 1349 RTHDNTADGR-----IVFMKEPKLAENLPQTP-----VCLAITLPRTYTG 1393
 QY 1111 PPEGVVSPDPHSSYSGLS-----GRPHFYNRAAVPATGSSLLIS-----MT 1155
 Db 1394 PEP--MWSKK--LFVSEASILEKYVGAFFETAPDMPLAHVALLIGADMHRLARALEV 1448

QY 1156 VDVCSEKLR-QIEGLDQNMMPQYCTTIKANINGRVLSCQNIDELKREMANFGDMHLR 1214
 Db 1449 PIDIRVNRHQVGLG-----AVTILRIWI-----FLKKQATPVA--LR 1486
 QY 1215 SVALEKRVESQVPEDEPRF--LNENSSAPVPH--GESARRSS-----HTELPL 1259
 Db 1487 SAL--QRIGRDDVREMDRAEKLDELGPVSHISGPSITLSLLEVAGDGRRRRAEVYM 1544
 QY 1260 TELS-SQPTYLNFSE-----ELNTIGLDE--GAPRHSNLSWISO 1297
 Db 1545 AOCORLAQEPFOOVYNGTTPGDPEBPKQSFHEEEVAVSIRTVTRERHVDSENGP 1604
 QY 1298 TRRTSLSLNSQDSIEISKLTDYV-----QAEY--RDAYREYIAQMSOLEGSTG 1346
 Db 1605 IVEERTITTTEDDAVANEELIVDKIVPLNEEQEKMRAVREMANFEQDETSEKGTIFG 1664
 QY 1347 SSTISGRSSPHSTYVYIGSSSGSISHLEDRGREGELKODGRKSLMRGVDYSS 1406
 Db 1665 CQT-----THEREKDDGSLKTTW--KDSHVRO-----IFPDGGETSANE 1703
 QY 1407 SGVSTNEASP-LDPTIETD 1424
 Db 1704 TGLSSGDAOTIMPTTKED 1722

RESULT 5

US-09-065-474-33

; Sequence 33, Application US/09065474

; Patent No. 6063599

; GENERAL INFORMATION:

; APPLICANT: Tang, Liang

; APPLICANT: Blehm, E. Scott

; TITLE OF INVENTION: DIOPHTERIA AND BRUGIA ANKYRIN

; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND

; NUMBER OF SEQUENCES: 171

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Carol Talkington Verser, Ph.D.

; STREET: 1825 Sharp Point Drive

; CITY: Fort Collins

; STATE: Colorado

; COUNTRY: USA

; ZIP: 80525

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: Windows 95

; SOFTWARE: Wordperfect for Windows, Version 7.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/065,474

; FILING DATE: 24-APR-1998

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Verser, Carol Talkington

; REGISTRATION NUMBER: 37,459

; REFERENCE/DOCKET NUMBER: HW-5-C1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 970/493-7272

; TELEFAX: 970/484-9505

; INFORMATION FOR SEQ. ID NO.: 33:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1745 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-09-065-474-33

Query Match 5.3%; Score 466.5; DB 3; Length 1745;
 Best Local Similarity 20.4%; Pred. No. 1e-23;
 Matches 358; Conservative 235; Mismatches 575; Indels 591; Gaps 73;

APPLICATION NUMBER: US/09/557,034
FILING DATE: 21-Apr-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/065,474
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Verser, Carol Talkington
REGISTRATION NUMBER: 37,459
REFERENCE/DOCKET NUMBER: HR-5-C1
TELEPHONE: 970/493-7272
TELEFAX: 970/484-9505
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 1745 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 33:
US-09-557-034-33

Query Match 5.38; Score 466.5; DB 4; Length 1745;
Best Local Similarity 20.4%; Pred. No. 1e-29;
Matches 358; Conservative 235; Mismatches 575; Indels 591; Gaps 73;

14 EENIPALKALECKVDNERGCGPTPLMLAEOGVVEIVKELINKGANCLEDIDN-- 71
207 KDDTKATLLDQNEHSVDTSKSGPTPLHIAHGENVAQDLLEGAVANVQARHNS 266
72 -----W-----TALSASKEGHIHYEELKSGASD--- 97
267 PLHATKMGRTNNVSLLAGAVIDCRTLPLPHKASRSGHDYVLLLEGAPISAK 326
98 -----EHRDMGWTALMW 110
327 TKNGLAPLHMAQVDVTVLPLHVAHAGHVRAKLLDRNAPRNALNGFTPLHI 386
111 ACYKGTVDYVELLSHGANSVTGLQSVPIIIMAGRGHADVHLLONGAVNCSDKY 170
387 ACKNRKIVELKTHIAIEAT-TEGSLSPHVAAMGAINIVTYLLDQGANADAVTR 445
171 GTPPLVAARKGHCLEKHLGADVDQGANSMALIVAAGGYTQSYKEILKRPNV 230
446 GEPPLHIAARANDIYRLVNRGAQYDAABELQPLHIASRLGNDIYILLQANASP 505
231 NLDDKQNTALMAASKGHEIYVODLDACTYVNIIPDRSGDYVIGAVRGHVEIYAL 290
506 NAATRDLYTPLHIAAKGQEVAAIIMDHGTDLTKKGTPLHIAAKYGNLPVAKSL 565
291 OKYADIDIRGODNKTALYMAVEKGNATMVARDILOCNPDEICTKDGFTPLIKATKMNIE 350
566 ERGPVVEIGKNQVTPHVAHAYNNNDKVALLLLENGASAHAAKNGTTPHIAAKNOMD 625
351 VVELLDKGAQVAVDKGDTPLHVALIRGRSRLAELLRN----- 391
626 IASTLLHYKANANASAKGFTPLHIAQEGHREMAALLLENGAKVGAQANGLPMLICA 685
392 -----PKDGLLIRPKAGETPYNIIDCSHOKSILTQIFGARHLSPTEDGMLG 440
686 QEDRVSAEELVVENAIDPKTAGYTPHVAHCH-----FG-----QINMVR 727
441 YDYSALADILSEPTQPPICVGLYQWQSGSKFLKTL-----EDEMKTFAQOQTEPL 495
728 FLIHGAKRVSITASTP-----LHQAAQGHNSVRYLLEHGASNVHTSTGO----- 777
496 FQSMILVFTLLLCGLGVFAFPVYDNLAIISLSFLAIYIFIVIFGGRREGESW 555
778 -----TPLSI----- 782
556 NMAMALSTRARHIGYELLFKLMFVNPPELPEOTTKALVRF-----LFTDYN 604
783 -----AERLGYSVVEALKTITETVITETTVTEERYKPPONPEAMNETMFSD-- 830

605 RLSSVGGENSLAEIATLSDACEREGFL--ATRLFVRVTEESGKK----- 650
831 ---SEDEGEDNQITANNAHADSESLTKGIMDSTGCHLIHATEPTLSRSEVEEGIDGDDA 888
651 -----KMKTCCLSPVIFLEI-----VGCIIAGIT 676
889 LIRKAQHEDITAMADPSLADSLPQNVITIMRTMQPSFLISPMVDARGAMRGCHSGVR 948
677 LAIFRVPDKHLTVNALISIASVYGLAVLNCRTWQVYDLSLN-----SQRRKLSAAS 732
949 IIPPRKAQOPTRV-----TCR--YLGKOKLADHPPLSGEAL--AS 986
733 KLHLKLSSEG--FMKYLKCEVELAMAMAKTIDSFQNOIRLVYII--DG-----LDAC 780
987 RILEMAPHGAKFLGVILEVPFASL-----GREREYITLSDGQMKHEQLBAT 1038
781 EODKVLQMLD-----TVRLFSKGPFIATFASDPHIIKAINQNS 822
1039 E-DAQVEVLNESFDEAEISQLDLHTSRITRLITNDPFW--YFA-----VTVRVRQEVHC 1090
823 VLKDSNINHDMRNIV-HLPVFLNSRGLSNARKLYSATNGDITCSDTGTGTQEDTBR 881
1091 VQPEGV-----ILSSVPHVQAIFFDGSJLTIKYVQA-----QP 1127
882 VSQNSLGEWTKL-GSKTALNRDQY--RRQMQRTITROMSPDLTK--LVTEDMFSDI 935
1128 VFOET--VTRHGNRVAVSPIYVPEPRKRRKHKTICILPQSSKMGMLTQYSGPQ 1184
936 SPQTRRLNIY-----SVTGRLLRANOITENMRSLASMINLQD-W----- 976
1185 EPTTLRLCSKTGGSSPAQWEDITG--TQULFTGEDVSPFTTVASRFLMDCQPTPD 1240
977 -----PYRSMILYIEE--EG-----LPDQTKLTM--YERISKNI 1010
1241 AARMAQEVYNEAIVAYAKFLI-FARRFPAGGLRLFCMDDBEDDTLEKQEPFI-I 1298
1011 PTKDVEPL-----LEIDGRNEFEVFLSSRTPLVARDYT--FLPCTVN--LDPKL 1059
1299 AKSKDEVLSGRHOFLEFSGNL-----LPITSSDQSLFTFLPQEKRLAFMKI 1348
1060 R-----ELIADYARAREQINIGLAVPLPLHGGPPRPSGYSOPASVCSA--SFNG 1110
1349 RHTDNETAADR-----IVFMKEPKLRAENLPPQF-----VCTLAITLPEYTG 1393
1111 PRPGGVSPQPHSSYSGIS-----GPHPTYNAAVATGSSILLSS-----MT 1155
1394 PEP--MVSKR-----LEFSEASLTERKYVGAFHETAEPDNPLAHVALIIGADWHRLARALEV 1448
1156 VDVCERKL-QLEGDQNMMPQYCTTIKANTNGRVLSCNIDELKKEAMNFGDMHLFR 1214
1449 PDIIDQVHQLVLE-----AVTILRIWI-----FLAKKQATVVA--LR 1486
1215 SMLERNVSQVPEDPFR--LNENSSAPVPH--GESARRSS-----HTELPL 1259
1487 SAL--ORIGRDVVRREMDRAEKLDGEGPVSHISGPISTLSTLLEVAAGDRRAHAYTM 1544
1260 TELS--SQPYTLNFSF-----ELNTLGLIDE--GAPRHSNLSMSQ 1297
1545 AQRALQEPFQOVGNGTPGDPEBKQSFHEEREVAVSSIRTVVTRTERIVHDSNGP 1604
1298 TRRTPLSLNSODSSIEISKLTDV-----QAEY--RDAREYIAQSOLEGGTG 1346
1605 IYERKITTITTEDVAVNEEIVDKIYPLNEEQRKMDRAMEVEMNPEQDTSKEGTFG 1664
1347 SSTIGRSSPHSTYYIGSSSGSISHTLEQGRGEGELKQDGRKFLMKRGDVIDYS 1406
1665 CQT-----THEKEDDQGSILKTM-----KDSHVRO-----IFPDGGETSANE 1703
1407 SGVSTNEASP-LDPTIRED 1424
1704 TGLSSGDAOTIMPTTKED 1722

RESULT 7
 US-09-082-059-2
 ; Sequence 2, Application US/09082059A
 ; Patent No. 6225086
 ; GENERAL INFORMATION:
 ; APPLICANT: Devorajan, Jon S.
 ; APPLICANT: Morrow, Prasad
 ; TITLE OF INVENTION: No. 6225086el Ankyrin Proteins and a Method for Their Identification
 ; FILE REFERENCE: 44574-5002-US
 ; CURRENT APPLICATION NUMBER: US/09/082,059A
 ; EARLIER FILING DATE: 1998-05-21
 ; EARLIER APPLICATION NUMBER: 60/047356
 ; EARLIER FILING DATE: 1997-05-21
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 2
 ; LENGTH: 1088
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-082-059-2

Query Match 5.1%; Score 449.5; DB 4; Length 1088;
 Best Local Similarity 31.1%; Pred. No. 1.2e-28;
 Matches 118; Conservative 74; Mismatches 184; Indels 3; Gaps 2;

QY 14 EEENIPALKLECKKDYDERNEGOTPLMLAEGNGVEIKELKNGANCNLEDLDMWT 73
 DB 30 KKNRIKVMELLKHGASIOAYTESGLTPIHVAAPFGHNVISQLMHGASPTTNVNGET 89
 QY 74 ALISAKSGHIIYVELLSGASLEHRDMGWTALMMACTYKRTDVEYELLSHGANP-SV 132
 DB 90 ALHMAARSQAENVVRYLVODGAQVEAKAKDQPTPLHISARLGKADIVQOLLQOGASPMVA 149
 QY 133 TGLQSVYPIIIMARSHADIVHILLONGAKVNSCDKGTPTVYMAAKGHECYKHLA 192
 DB 150 TTSGTT--PLHLSAREGHEDVAAPFLLDHGASISTITTKGFPPLHVAAYGKLEVANILLQ 207
 QY 193 MGADYDOEGANSMTALIVAVKGYQSVKEILKRNPNVNLTDKDGNTALMAKSGHEI 252
 DB 208 KASPDAGKSGLTPLHVAHNDQKVALLLDQASPHAAKKNYTPPLHIAAKKNQDI 267
 QY 253 VODLDAGTYNIPRSGDVTILGAVRGHVEIYRALLQYADIDIRGDNKTAIYMAVE 312
 DB 268 ATTLEYGADANAVYRGIASVHLAAQEGHVDVSLIGRANAVLSNKSGLTPLHLAAQ 327
 QY 313 KGNATMVRDILQCNPDTEICTKDETPLIKATKMNIEVVELLDKGAASVADKGDTP 372
 DB 328 EDRVAVAEVLVNOGAHVAQTKMGYTPPLHVGCHGNIKIVNELLOHSAKAVNAKTNGTTP 387
 QY 373 LHVAVRGRSRLAEILLRN 391
 DB 388 LHOAAQGHHTHINVLQN 406

RESULT 8
 US-08-810-712-10
 ; Sequence 10, Application US/08810712G
 ; Patent No. 6160106
 ; GENERAL INFORMATION:
 ; APPLICANT: Yeda Research and Development Co. LTD
 ; TITLE OF INVENTION: Tumor Suppressor Genes, Proteins Encoded Thereby and
 ; FILE REFERENCE: sequencelist
 ; CURRENT APPLICATION NUMBER: US/08/810,712G
 ; CURRENT FILING DATE: 1997-03-03
 ; EARLIER APPLICATION NUMBER: PCT/US94/11598
 ; EARLIER FILING DATE: 1994-10-12
 ; NUMBER OF SEQ ID NOS: 31
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 10
 ; LENGTH: 1423
 ; TYPE: PRT

ORGANISM: Homo sapiens
 US-08-810-712-10

Query Match 4.7%; Score 414.5; DB 4; Length 1423;
 Best Local Similarity 21.6%; Pred. No. 1.9e-25;
 Matches 221; Conservative 156; Mismatches 335; Indels 309; Gaps 36;

QY 5 ISQSVINVEEENIPALKLECK--DYDERNEGOTPLMLAEGNGVEIKELKNGA 62
 DB 336 VMALHAIINDVDPGLHLSGLSNVDNQPKHTPPLLLIAGCGNIQLILKRG 395
 QY 63 NCNLEDNMTALISAKSGHIIYVELLSGASLEHRDMGWTALMMACTYKRTDVEY 122
 DB 396 RIVQDKGSNAYVMAARHGHVDTPLFSLNKPRLPVYKSGEMALHVAARGHADVA 455
 QY 123 LSHGANPSTVGTQYVYPIIIMARSHADIVHILLONGAKVNSCDKGTPTVYMAAR 182
 DB 456 TCASAOIPISRTKEETPLHCAAMHGYSVAKALCEACGNVINKNREGETPLTFAS 515
 QY 183 HLECYKHLAMGADVDOEGANSMTALIVAVKGYQSVKEILKRNPNVNLTDKDGNTAL 242
 DB 516 YHDIVECLAHGAD-----LANCKDGHIAH 542
 QY 243 IASKEGHIETVODLDAGTYNIPRSGDVTILGAVRGHVEIYRALLQYADIDIRG 302
 DB 543 LAVRCQMEVITKLLSGCFVDYQDRHNGTPLHVACKDGNMPLVALCE----- 591
 QY 303 NKTALVAVKGNATVYRDLQCNPDTEICTKGEPTPLKATMRNIEVVELLDKGA 362
 DB 592 -----ANCNLD--ISNKGRTPLHAAANGILDVRYCLMGASV 629
 QY 363 SAVDKGDTPLHVAIRGRSHRLAELLRNPCKDRLLYRNKAGETPNIDCSHOKSIL 422
 DB 630 EALTQCKTMDLARBQHEHVAGLARLKD-----THRGFLIQ 670
 QY 423 IFGARHLSPTETDGMIGDLYSSALADILSEPMPPICVGLYAOVSGKSFLLKIED 482
 DB 671 -----LRPQ-----NLQPRIKLFGHSGKGTIVESTK- 701
 QY 483 EMKTFAGQCEPLPQFSWLVFLTLGGLGLVF-----AFPDNTNLA 526
 DB 702 -----CGLSFFRRRRPRSSINSSRPPSPPLASKET 734
 QY 527 IATSLFALITYFFVIYFGRRBESNMWAMALSTRARHIGYELLFKMFVNPRL 586
 DB 735 VSVSINN-----YPCGENSVSRSMFEPGLK--GMLE-----VFAPPIH 776
 QY 587 P-----EQTAL-----PYRPLFTDYNRLSVSGETSLAEMTA 620
 DB 777 PHCSADDOSTKAIIDIONAYLNGVDFSVWFSGNPFYFCYDY--FAANDPTSHVVF 833
 QY 621 TLDACERERG--PLATRLFRVFRTEES--QGKKMKKTCTCLPSFVFLFYGGI 674
 DB 834 SLEPFEIQLNPVIFMLSFSLKSLVPEEPLARGKLL-----NPLQVVL----- 877
 QY 675 ITLLAIFRVDKHLVYNALISASVYGLAFVLCNRTMWOVLDLSLNSQKR-----LHSA 730
 DB 878 -----VAHADIIMVPRPAGGEFGDKT-----SLKEIRNFGDLH-I 917
 QY 731 ASKLHLKSGEFG--MKVLKCEVELMAAMAKTIDSTFONQRLVYIIDGLACEDKYL 786
 DB 918 SNKLFLVDGASGSKDKVLRNLIQ--EIRSOIVSVCPPMHL-----CE--KII 963
 QY 787 QMDTVRVLFESKPFIAI--FASDPHIIKAIQNLNLSLRDINSNGHDMRNIYHLPE 844
 DB 964 STLPWRKLGNQMLSDQFYD-----VODQINPLASED-----LRIHQO-- 1007
 QY 845 LNSRGLSNARKPLVTSATNGDITCSDPTQEDTDPRVSQNSIGENTKIGSKTALNR-RD 903
 DB 1008 LHSSTGEIN--IMQSEYQVDVLLD-----PRWLCTNVLGKILSTFPRALAHNRG 1055
 QY 904 TYRRQMQRTITRQMSFDITKLLVYEDWFS-DISQPMRRLINIVSVTGRLIRANOITFN 962

Db 1056 RTVEDIQRIVPDSVDELLQILDAMDICARDLSSGTM-----VDYPA-LIKTDNLHRS 1108
QY 963 W 963
Db 1109 W 1109

RESULT 9
US-09-196-387-2

Sequence 2, Application US/09196387
Patent No. 6277613
GENERAL INFORMATION:
APPLICANT: de Lange, Titia
APPLICANT: Smith, Susan
TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS
TITLE OF INVENTION: OF USE THEREOF
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue, 4th Floor
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/196,387
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/095,225
FILING DATE: June 10, 1998
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-230 CIP1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1327 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
US-09-196-387-2

Query Match 4.5%; Score 395.5; DB 4; Length 1327;
Best Local Similarity 23.3%; Pred. No. 6,9e-24;
Matches 167; Conservative 80; Mismatches 211; Indels 259; Gaps 17;

QY 14 EENIRPAKALKKCKDVERNECGOTPLMLAEGQVETIKLNKANGCNLEDDNMT 73
Db 348 EERLMALPLNCHASDGRK---STPLHLAAGYNRVRIVOLLHGADVHAKDKGLV 404
QY 74 ALISASKEGHIHVEILLKSGASLEHRDMGWTALMWACYKGTDVVELLSHGANSV- 132
Db 405 PLNMASCYGHEYTELLKKGACVAMADMQFTPLHDAASKNVEVCSLLSHGADPTLV 464
QY 133 -----T 133
Db 465 NCHGSAVDMAPTPPELRELTVEFKGSHLQAAREADLAKVKTALLETINRKQPSHET 524
QY 134 GLQSYV-----YPIIWAAGRHADIVHLLQNGARV 164
Db 525 ALHCAVASLHPKRYQTELLRKGANVNEKNKDFMTPLHVAERANDVMEVYLHKGAKM 584

QY 165 NCSDRTGTPPLVAAKRGHECVKHLAMGADVDCGANSMTA----- 207
Db 585 NALDTLGGTALHRAALAGHLQTCRLLLSGSPDSIISLGFTPAQNGNVAQOILSESP 644
QY 208 -----LIVAVKGYTQSVKEILLKRNPNVLTDKDG--NTALMAKKEGHIETVDL 256
Db 645 IRTSDVDYRLLESKAGDLETVKOLCS--SONVNCRLDERHSPHLFAAGYNRVSVVEYL 703
QY 257 LDAGTYVNIPIRSGDVLIGAVNGHVEIVRALLQKYADIRGQDNKALYAVEK-- 314
Db 704 LHHGADVHAKDKGGLVPLNMASCYGHEYTEALLVRCAGAVNVADLKKFTPLHDAAKGY 763
QY 315 -----NATVWDIIO-----C 325
Db 764 EICKLLKHGADPTKKNRDNTPDLVKEGDDIDQLLKGDAALLDAAKGGLARYOKC 823
QY 326 NPDEICTKD---GFTPLKATKRNIEVEVLLDKGAVSAVDKKGDTPLH----- 374
Db 824 TPEINCRDPTQGNSTPLHLAAGYNMLEVAEYLLERGADVMAQDKGLPLHMAASYGHV 883
QY 375 -----VAIRGSRRLAEILLRNPDKGRLLYPRNKAETP 408
Db 884 DIAALLIKYNTCVNATDKKAFTPPLHDAOKGT--QCALLLHAGADPTM---KNQGGTP 939
QY 409 YN-----IDCSHOKSIIT-----QIFGARHLSPTETDGMGLDYSSALADI 451
Db 940 LDATADDIRALLIDAMPPEALPTCFKPOATVVASLISPASTPS-----CLSAASIDN 994
QY 452 LSEPTMQPIICVGLYQWQSG-----KSFLLKLEDEKMT 486
Db 995 LIGPLAE--LAVGASNAGDGAAGTERKEGEVAGLDMNISQPLKSLGLEHLNDIET 1049

RESULT 10
US-09-196-387-10

Sequence 10, Application US/09196387
Patent No. 6277613
GENERAL INFORMATION:
APPLICANT: de Lange, Titia
APPLICANT: Smith, Susan
TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS
TITLE OF INVENTION: OF USE THEREOF
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue, 4th Floor
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/196,387
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/095,225
FILING DATE: June 10, 1998
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-230 CIP1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:

```

;      LENGTH: 949 amino acids
;      TYPE: amino acid
;      TOPOLOGY: linear
;      MOLECULE TYPE: protein
;
US-09-196-387-10

```

Query Match	4.48;	Score 387.5;	DB 4,	Length 949;
Best Local Similarity	25.28;	Pred. No. 1.8e-23;		
Matches 144;	Conservative 59;	Mismatches 177;	Indels 191;	Gaps 11.

```

QY      14 EEEINIPALFALLLECKKDDVERNECQOTPLMLAAEGNEIYKELLKNGANCNIEDDNNM 73
      348 EEEMLALLPLPWNCHASDGR-----STPLHLAAGNRIYQJLLQJGADVHAKDKGVL 404
Dp
QY      74 ALISASKEGHIIHVEBLKSGASLEHRDMGCGTALMMACYGRDYYVELLSHGANSV- 132
      405 PLHNACSYGHEYETELLKHGACVAMDLMOFTPLHEAASKNRVECSLLLSHGADPTLV 464
Dp
QY      133 -----T 133
Dp      465 NCHGKSAVDMAPTPELRERLYEFKGHSLQAREADLAKYKKTLEIINFKOPOSHET 524
      134 GLQSY-----YPLIWAAGRGHADIVHLLLQNGAKV 164
      525 ALHCAVASLHPKKQYOTELLKKGANVENKNDFTPLHVAEEAANDVMEVLHKHAKM 584
Dp
QY      165 NCSDKYGTTPLYWAARKGHLEVCYKLLAMGADVDOEGANSMTA----- 207
      585 NALDTLGGYALHHAALAGHLQTCRLLLSGSDPSITSLQGTAAQMGNEAVQJLSESTP 644
Dp
QY      208 -----LIYAVKGYQTSYKELLKRPNNPLTDKGG-NTALMTASKEGHEIYQDL 256
      645 IRTSDVYRLLEKSKAGDELYKQCS-SQNNNCDELRGSRSTPLHFAAGYNRVSVEYL 703
Dp
QY      257 LDAGTYVNIIPDRSGDVTLLIGAVRGGEIVETRALQYADIDIRQDNKTALTYAWEKG- 314
      704 LHHGADVHAKDGLVPLHNACSYGHEYAEELVHNGASVAVADLMKFTPLHEAALKGX 763
Dp
QY      315 -----NATVRDILQ-----C 325
      764 EICKLLKHGADPTKKRNDGNTPPLDLYKRGDTHQDLKDBAALLDAAKKGLARVQKLC 823
Dp
QY      326 NPOTELCTKD---GELPPLKATKMRNIEVELLDKGAKVASVDKGGTPLHVALIRGRSR 382
      824 TPERINCRDTQGRNSTPLRLHLAGYNNLEVAEYELLEHGADVAADKGGIPLHNA----- 877
Dp
QY      383 RLAEILLRNPKDGRLLLYRPNKAGEPNYIDC 413
Dp      878 -----ASTGGCLARVQKLC-TPERINCC 898

```

US-09-031-485-28
 SUDLT 11
 Sequence 28, Application US/09031485
 Patent No. 5824306
 GENERAL INFORMATION:
 APPLICANT: Tang, Liang
 APPLICANT: Blehm, E. Scott
 TITLE OF INVENTION: PROTOFILARIA AND BRUGIA
 TITLE OF INVENTION: PROTEINS, NUCLEIC ACID
 TITLE OF INVENTION: USES THEREOF
 NUMBER OF SEQUENCES: 85
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Carol Talkington Verser, Ph.D
 ADDRESSEE: Hesk Corporation
 STREET: 1825 Sharp Point Drive
 CITY: Fort Collins
 STATE: Colorado
 COUNTRY: USA
 ZIP: 80525
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible

```

OPERATING SYSTEM: Windows 95
SOFTWARE: Wordperfect for Windows, Version 7.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/031,485
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/847,429
FILING DATE: 24-Apr-1997
ATTORNEY/AGENT INFORMATION:
NAME: Verser, Carol Talkington
REGISTRATION NUMBER: 37,459
REFERENCE/DOCKET NUMBER: HM-5
TELECOMMUNICATION INFORMATION:
TELEPHONE: 970/493-7272
TELEFAX: 970/484-9505
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 348 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-031,485-28

```

query Match	4.28;	Score 376;	DB 2;	Length 348;
Best Local Similarity	34.08;	Pred. No. 2.5e-23;		
Matches 113;	Conservative 47;	Mismatches 134;	Indels 38;	Gaps 4

```

OY      45  AAGQGVVEIYKLEIKKNGANCNEDEDDNTTALISAKSECHIIYVELLSGSGSYENRDMG  104
Dd      44  AARAGMDRLVELLSRGSTDFINFCNMNGINALHLIAKSEGHNEVEVELLKRADYDAATRK  103
OY      105  WTALMAACKGRDVEYELLSHGANCPTYGQIYVYPTIYMAAGGHADYHILLIIONGAKV  164
Dd      104  NTLHLIASLAGOELLITVYLVENGANVVOISL-GETPIYMAQENHESVRYILLAHNAQ  162
OY      165  NCSDKYGTTPYLYMAARKGHLEVCYKILLMAGADVDOEGANSMTALIVAKGQYQSVKEIL  224
Dd      163  ALSTEDGEPFLAVALAQGHDRVAAVLL-----ENDTRGKVRRLPALHIAAKDDTFAAAILL  218
OY      225  KRPNPNVLDKOCNTMTAMTAKSEGHLEIYODLIDAGTVYVNIIPDRSGDPTVLIGARGGHE  284
Dd      219  QNEHNSVSVKSGFTPLHTAHYGNENNAQILLLEGANVNTQAKHINSPLHVALTKWRTN  278
OY      285  IVRALLQYADIDIRGQDNKTALUYAVEKGNATVYRDLQCNPDTEICTKDGDETPILKAT  344
Dd      279  MWSLLAHGAVIDCR-----TRDLL-----TPHLCAS  305
OY      345  KMRIEYVELLDLKGKASVADKGGTPYHYA  376
Dd      306  RSGHDQVVDLLEKGAIPISAKTRKNGLAPLHMA  337

```

RESULT 12
 US-08-847-429A-28
 : Sequence 28, Application US/08847429A
 : Patent No. 5827692
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: Tang, Liang
 :
 : TITLE OF INVENTION: DIBOPIRLARIA AND BRUGIA ANKYRIN
 :
 : TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
 :
 : TITLE OF INVENTION: USES THEREOF
 :
 : NUMBER OF SEQUENCES: 85
 :
 : CORRESPONDENCE ADDRESS:
 :
 : ADDRESSEE: Carol Talkington Verser, Ph.D.
 :
 : ADDRESSEE: Heska Corporation
 :
 : STREET: 1825 Sharp Point Drive
 :
 : CITY: Fort Collins
 :
 : STATE: Colorado
 :
 : COUNTRY: USA
 :
 : ZIP: 80525
 :
 : COMPUTER READABLE FORM:

```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: WordPerfect for Windows, Version 7.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/847.429A
FILING DATE: 24-APR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Verser, Carol Talkington
REGISTRATION NUMBER: 37,459
REFERENCE/DOCKET NUMBER: HW-5
TELECOMMUNICATION INFORMATION:
TELEPHONE: 970/493-7272
TELEFAX: 970/484-9505
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 348 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-847-429A-28

```

```

Query Match          4.2%  Score 376; DB 2; Length 348;
Best Local Similarity 34.0%; Pred. No. 2.5e-23;
Matches 113; Conservative 47; Mismatches 134; Indels 38; Gaps 4;

```

```

QY 45 AAEQGNVEIYKELLKNGANCLELDNMTALISASKEGHIIVEELLKSGASLEHRDMG 104
DB 44 AARGNDRIYVELLRSGDITNCNANGLNLHLSKEGHEVRELLKRRKADVDAAATRG 103
QY 105 WTALMAYCYGRDVTVELLSHGANSVYGLQSVYPIIYMAAGRGADIVHLLONGAKV 164
DB 104 NTLALHSLAGOEIYTVLVENGANVVOGLN-GFTPLVMAQENHESVYRLLANANQ 162
QY 165 NCSKRYGTPPLVMAARKGLECYKHLAMGADVQEGANSMTALIVAVGQTSYKEIL 224
DB 163 ALSTEDGTPPLVAVLQGHDRVAVLL-----ENDTRGXVRLPALHIAKKDDTKAATLL 218
QY 225 KRPNVNLTDKQNTALMTASKEGHIETVQDLIDAGTYVNIIPDRSGDTVLIGAVRGHVE 284
DB 219 QNEHNSDVTSKSGFTPLHIAHYNENVAQLLEKGANVYQARHNISPLHATKMGRTN 278
QY 285 IYRALLQYADIDIRQDNKTALYMAVEKGNATVNDIQCNDPEICTKDETPLIKAT 344
DB 279 MVSLLAHGAVIDR-----TRDL-----TPLCAS 305
QY 345 KMRNIEVELLDKGAQVSAVDKKGDTPLHVA 376
DB 306 RSGHDQYVDLLEKGAQISAKTKNGIAPLHMA 337

```

```

RESULT 13
US-09-065-474-28
Sequence 28, Application US/09065474
Patent No. 6063599
GENERAL INFORMATION:
APPLICANT: Tang, Liang
APPLICANT: Blehm, E. Scot
TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
NUMBER OF SEQUENCES: 171
CORRESPONDENCE ADDRESS:
ADDRESSEE: Carol Talkington Verser, Ph.D.
ADDRESSEE: Heska Corporation
STREET: 1825 Sharp Point Drive
CITY: Fort Collins
STATE: Colorado
COUNTRY: USA
ZIP: 80525
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

```

```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: WordPerfect for Windows, Version 7.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/065.474
FILING DATE: 24-APR-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Verser, Carol Talkington
REGISTRATION NUMBER: 37,459
REFERENCE/DOCKET NUMBER: HW-5-C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 970/493-7272
TELEFAX: 970/484-9505
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 348 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-065-474-28

```

```

Query Match          4.2%  Score 376; DB 3; Length 348;
Best Local Similarity 34.0%; Pred. No. 2.5e-23;
Matches 113; Conservative 47; Mismatches 134; Indels 38; Gaps 4;

```

```

QY 45 AAEQGNVEIYKELLKNGANCLELDNMTALISASKEGHIIVEELLKSGASLEHRDMG 104
DB 44 AARGNDRIYVELLRSGDITNCNANGLNLHLSKEGHEVRELLKRRKADVDAAATRG 103
QY 105 WTALMAYCYGRDVTVELLSHGANSVYGLQSVYPIIYMAAGRGADIVHLLONGAKV 164
DB 104 NTLALHSLAGOEIYTVLVENGANVVOGLN-GFTPLVMAQENHESVYRLLANANQ 162
QY 165 NCSKRYGTPPLVMAARKGLECYKHLAMGADVQEGANSMTALIVAVGQTSYKEIL 224
DB 163 ALSTEDGTPPLVAVLQGHDRVAVLL-----ENDTRGXVRLPALHIAKKDDTKAATLL 218
QY 225 KRPNVNLTDKQNTALMTASKEGHIETVQDLIDAGTYVNIIPDRSGDTVLIGAVRGHVE 284
DB 219 QNEHNSDVTSKSGFTPLHIAHYNENVAQLLEKGANVYQARHNISPLHATKMGRTN 278
QY 285 IYRALLQYADIDIRQDNKTALYMAVEKGNATVNDIQCNDPEICTKDETPLIKAT 344
DB 279 MVSLLAHGAVIDR-----TRDL-----TPLCAS 305
QY 345 KMRNIEVELLDKGAQVSAVDKKGDTPLHVA 376
DB 306 RSGHDQYVDLLEKGAQISAKTKNGIAPLHMA 337

```

```

RESULT 14
US-09-557-034-28
Sequence 28, Application US/09557034
Patent No. 6365569
GENERAL INFORMATION:
APPLICANT: Tang, Liang
APPLICANT: Blehm, E. Scot
TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
NUMBER OF SEQUENCES: 171
CORRESPONDENCE ADDRESS:
ADDRESSEE: Carol Talkington Verser, Ph.D.
ADDRESSEE: Heska Corporation
STREET: 1825 Sharp Point Drive
CITY: Fort Collins
STATE: Colorado
COUNTRY: USA
ZIP: 80525
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

```

OPERATING SYSTEM: Windows 95
 SOFTWARE: Wordperfect for Windows, Version 7.0
 CURRENT APPLICATION DATA: US/09/557,034
 APPLICATION NUMBER: US/09/557,034
 FILING DATE: 21-Apr-2000
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 09/065,474
 FILING DATE: <Unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Verser, Carol Talkington
 REGISTRATION NUMBER: 37,459
 REFERENCE/DOCKET NUMBER: HW-5-C1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 970/493-7272
 TELEFAX: 970/484-9505
 INFORMATION FOR SEQ ID NO: 28:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 348 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 28:
 S-09-557-034-28

[illegible]

```

? FEATURE:
? NAME/KEY: misc_feature
? OTHER INFORMATION: xaa is any amino acid
? NAME/KEY: misc_feature
? LOCATION: (1102)..(1102)
? OTHER INFORMATION: n is any nucleic acid
? NAME/KEY: misc_feature
? LOCATION: (2650)..(2650)
? OTHER INFORMATION: n is any nucleic acid
US-09-350-982C-5

```

Query Match	4.2%	Score	373;	DB	4;	Length	1166;
Best Local Similarity	23.4%	Pred. No.	4.5e-22;				
Matches	144;	Conservative	62;	Mismatches	181;	Indels	228; Gaps
							12;
QY	19	PAIKALLEKCD	-----	VDREBCCQ	-TPLMLAABEGNVEIYKELLKNGA	62	
DB	23	PAARELFECACRGNDYERKRLVTPEKVNSRDAKRSKTPHLHLAGFRDVEYLLQNGA	82				
QY	63	NCLBEDLDWTLTASKEGHHIYEELKSGALSERDMGWTALTMACSYGRDVEL	122				
DB	83	NWQARDGGCLPLHNAACSFGEAEVYNLLRRGADPNRMWNTPLHEAIAIKGIDVCIV	142				
QY	123	LLSHGANPSY	-----	TGLQY	137		
DB	143	LLOHGAEPTIRMTDGTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMAALLPLPNV	202				
QY	138	SVY-----PLIMAGRGHADIVHLLLONGAVNCSDDYGTGPLVMAARKGHE	185				
DB	203	NCHASDGKRSKTPHLHLAGTNRKVIYOLLQHDADVHAKDGDPLPLHNAACSYGHEVYTEL	262				
QY	186	-----CV-----	187				
DB	263	LYKHGACVNAMDLMOFTPLHEAASKRNVESCILLSYGADPTLLNCHNKSALIDLAPTPOL	322				
QY	188	-----KH-----	189				
DB	323	KERLAEFGHSHLLQAREADVTRIKKLLSLEWNEFKHPOTHETAXHCAAASTPYPRKROI	382				
QY	190	---LLMAGDADVDGEGANSTALIVAKGYTSQYKEHLKRPNVNLTDKDGNTALMIASK	246				
DB	383	CELLLRGAXINKTEFETPLHVASSEKANDXEYVVYKHEAKVNALDNGOTSLHRAAY	442				
QY	247	EGHI-----	278				
DB	443	CGHLOTORCLLSYGCDPNILLSQGTALOMGENNYOQLLOEG--TSLGNSBADROLLEBA	500				
QY	279	RGHVEIVRAL--LQKYTADIDIRGODNKNTALYMAVEKGNATMYBDILQCPDTEICTKG	336				
DB	501	KAGVEYEVKKLTGVQVNCRDIEGNS--TPLHFAAGYRNVSVVEYLLQHDADVHAKDKXX	559				
QY	337	ETPLIKATKRNIEVEYLLLDGAKYSAVDKGGDTPLHVALIRGSRKLAELLRNPKDGR	396				
DB	560	LVPRLHNAACSGHAEVEYELLVKHGAVVYVADLMKFTPLHEAIAAKGYEICKLILLQHDGAPT	619				
QY	397	LTVPRNKAGETPYNI	411				
DB	620	---KKNNDGNTPLDL	631				

```
Search completed: July 1, 2003, 14:40:44
Job time : 29.5 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2003, 14:34:04 ; Search time 89 Seconds
(without alignments)
3970.458 Million cell updates/sec

Title: US-10-021-571-4
Perfect score: 8884
Sequence: 1 MSYLISQSYINVEENIPALAKLLEKCDVDERNECGGTPMLIAEGNGNIEIVELIKN 1715

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_plant:*
10: sp_rodent:*
11: sp_virus:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_dactariap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8830	99.4	1777	4	Q9ULH0
2	8362.5	94.1	1762	11	Q9EGG6
3	8266	93.0	1715	11	Q9ERD4
4	5205.5	58.6	1031	4	Q9UFA2
5	2791	31.4	543	4	Q9H9E4
6	2379.5	26.8	1498	5	Q9W211
7	1832.5	20.6	1398	5	Q20109
8	1255	14.1	246	4	Q9NT37
9	666	7.5	129	4	Q9H889
10	557.5	6.3	1856	4	Q99407
11	544	6.1	426	16	Q8YTG9
12	523	5.9	1943	11	Q61307
13	522.5	5.9	1726	11	Q8VC68
14	522	5.9	1719	4	Q13768
15	521	5.9	4377	4	Q12955
16	519.5	5.8	2443	5	Q9VSA2

17	517	5.8	2622	11	Q70511	070511	rattus norv
18	506	5.7	1848	11	Q61302	061302	mus musculu
19	500.5	5.6	6994	5	Q17343	Q17343	caenorhabdi
20	498.5	5.6	1136	6	Q9N180	Q9N180	bos taurus
21	495.5	5.6	1159	5	Q9NCP8	Q9NCP8	drosophila
22	493.5	5.6	843	11	P97582	P97582	rattus norv
23	488.5	5.5	1486	4	Q8TEF1	Q8TEF1	homo sapien
24	485.5	5.5	1009	5	Q8SWY2	Q8SWY2	drosophila
25	482.5	5.4	784	5	Q9YCA7	Q9YCA7	drosophila
26	482.5	5.4	4001	5	Q8WRQ7	Q8WRQ7	drosophila
27	481.5	5.4	833	4	Q96186	Q96186	homo sapien
28	478.5	5.4	1549	5	Q9V4B1	Q9V4B1	drosophila
29	475.5	5.4	1549	5	Q24241	Q24241	drosophila
30	474.5	5.3	1599	11	Q99NH0	Q99NH0	mus musculu
31	474	5.3	1762	11	Q88521	Q88521	rattus norv
32	466.5	5.3	1188	4	Q9H288	Q9H288	homo sapien
33	463	5.2	1282	5	Q8T4F8	Q8T4F8	drosophila
34	463	5.2	2119	5	Q9VAU5	Q9VAU5	drosophila
35	463	5.2	2119	5	Q8T9G9	Q8T9G9	drosophila
36	451	5.1	1867	5	Q17486	Q17486	caenorhabdi
37	447.5	5.0	1815	5	Q17488	Q17488	caenorhabdi
38	447	5.0	460	4	Q9H6J9	Q9H6J9	homo sapien
39	447	5.0	743	4	Q8T846	Q8T846	homo sapien
40	442.5	5.0	1809	5	Q17487	Q17487	caenorhabdi
41	442	5.0	2039	5	Q17489	Q17489	caenorhabdi
42	441	5.0	627	4	Q96G77	Q96G77	homo sapien
43	438	4.9	686	4	Q9NXP0	Q9NXP0	homo sapien
44	438	4.9	1088	4	Q13484	Q13484	homo sapien
45	437	4.9	532	4	Q9H2U0	Q9H2U0	homo sapien

ALIGNMENTS

RESULT 1
Q9ULH0 PRELIMINARY; PRT; 1777 AA.
ID Q9ULH0
AC Q9ULH0;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE KIAA1250 protein (Fragment).
GN KIAA1250.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20039619; PubMed=10574462;
RA Nagase T., Ishikawa K., Kikuno R., Hirotsawa M., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XV.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 6:337-345(1999).
DR EMBL: AB033076; BAA86564.2; -.
DR InterPro: IPR002110; ANK.
DR Pfam: PF00023; ANK; 12.
DR PRINTS: PR01415; ANKYRIN.
DR SMART: SM00248; ANK; 11.
DR PROSITE: PS50088; ANK_REPEAT; 10.
DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
KW ANK repeat; Repeat.
FT NON_TER
SQ
SEQUENCE 1777 AA; 197209 MW; B6505923FBA5F143 CRC64;
Query Match 99.4%; Score 8830; DB 4; Length 1777;
Best Local Similarity 96.7%; Pred. No. 0;
Matches 1714; Conservative 0; Mismatches 0; Indels 58; Gaps 2;
1 MSYLISQSYINVEENIPALAKLLEKCDVDERNECGGTPMLIAEGNGNIEIVELIKN 60
|||||

Db 7 NSVLISQSVINVEEBENTPALKLEKCKDVERRECCGPTLMIABEONLEIYVELIKN 66
 QY 61 GANCNLEEDNDNTALISASKEGHVHIVEELLKGVNLEHRDMGWTALMAMCYKRTDYY 120
 Db 67 GANCNLEEDNDNTALISASKEGHVHIVEELLKGVNLEHRDMGWTALMAMCYKRTDYY 126
 QY 121 ELLLSHGANSVTGLOYSYPTIIMAGRGHADIYVLLLONAKVCSPKGYGTPLYMAAR 180
 Db 127 ELLLSHGANSVTGL-YSVPTIIMAGRGHADIYVLLLONAKVCSPKGYGTPLYMAAR 185
 QY 181 KGHLECVKHLAMGADVDOEGANSMTALIVAKGYTOSYKELIKRNVNLTDKGNTA 240
 Db 186 KGHLECVKHLAMGADVDOEGANSMTALIVAKGYTOSYKELIKRNVNLTDKGNTA 245
 QY 241 LMIASKEGHTETVOLLNAGTYVNIIPDRSGDYLIGAVRGHVEYVRLLOKXADIDIRG 300
 Db 246 LMIASKEGHTETVOLLNAGTYVNIIPDRSGDYLIGAVRGHVEYVRLLOKXADIDIRG 305
 QY 301 QDNKTALYWAWEKGNATWVRLLOCNPTETICTKGETPLIKATKRNIEVVELLDKGA 360
 Db 306 QDNKTALYWAWEKGNATWVRLLOCNPTETICTKGETPLIKATKRNIEVVELLDKGA 365
 QY 361 KYSAYDKGDPPLHAIKGRSRKLAELLRNPKGRLLYRPNKAGETPYNIDCSHOKSIL 420
 Db 366 KYSAYDKGDPPLHAIKGRSRKLAELLRNPKGRLLYRPNKAGETPYNIDCSHOKSIL 425
 QY 421 TQIFGARHLSPETEGMDLGYDLYSSALADILSEPTMPPICVGLYAOWGSKSLKLKL 480
 Db 426 TQIFGARHLSPETEGMDLGYDLYSSALADILSEPTMPPICVGLYAOWGSKSLKLKL 485
 QY 481 EDEMTFAGQOIEPLFOFSWMLVFLTLGGLGGLFAFVHPNGLIVSFLALXIF 540
 Db 486 EDEMTFAGQOIEPLFOFSWMLVFLTLGGLGGLFAFVHPNGLIVSFLALXIF 545
 QY 541 FVIVYFGGRREBSNNMAMVSTRLAKHIGYELLLKLMFNVPPLPQOTTKALPVPLF 600
 Db 546 FVIVYFGGRREBSNNMAMVSTRLAKHIGYELLLKLMFNVPPLPQOTTKALPVPLF 605
 QY 601 TDYNNLSVSGETSLAEMIALTSDACEREFGLATRLFRVFKTEPTQCKKKMKTKCCLPS 660
 Db 606 TDYNNLSVSGETSLAEMIALTSDACEREFGLATRLFRVFKTEPTQCKKKMKTKCCLPS 665
 QY 661 FVIFLFIIGCIISGITLLAIFRVDPKHLTVNAVLISASVGLAFVLCRTMWOVLDSL 720
 Db 666 FVIFLFIIGCIISGITLLAIFRVDPKHLTVNAVLISASVGLAFVLCRTMWOVLDSL 725
 QY 721 NSQKRLHNAASKLHKLKSEGFMKVLCCEVELMAMAKTIDSFQONOTRLVVIIDGLDAC 780
 Db 726 NSQKRLHNAASKLHKLKSEGFMKVLCCEVELMAMAKTIDSFQONOTRLVVIIDGLDAC 785
 QY 781 EODKVLQMDIVRVLFSGPFIAFASDPHIIKAINONLSVLDNSINMGHDYMRNVH 840
 Db 786 EODKVLQMDIVRVLFSGPFIAFASDPHIIKAINONLSVLDNSINMGHDYMRNVH 845
 QY 841 LPVFLNSGSLNARKFLVTSATNGDVPSCDTTGOEDADRVYSONSLGEMTKLSKTAALN 900
 Db 846 LPVFLNSGSLNARKFLVTSATNGDVPSCDTTGOEDADRVYSONSLGEMTKLSKTAALN 905
 QY 901 RDTYRRROMORTITROMSPDLTKLVTEDWFSOISPOTRMLINIVSVTRRLRANOIS 960
 Db 906 RDTYRRROMORTITROMSPDLTKLVTEDWFSOISPOTRMLINIVSVTRRLRANOIS 965
 QY 961 FNMDBLASMINLTEOWPYRTSMLILYEETEGIPDOMLKTIERISKINPTTADVPPL 1020
 Db 966 FNMDBLASMINLTEOWPYRTSMLILYEETEGIPDOMLKTIERISKINPTTADVPPL 1025
 QY 1021 EIDGDIRNEFVFLSRPVLVADVKYFLPCTVNLDPKRLIADVRAARQOISIGLAY 1080
 Db 1026 EIDGDIRNEFVFLSRPVLVADVKYFLPCTVNLDPKRLIADVRAARQOISIGLAY 1085
 QY 1081 PPLPLHEGPPRAPSGSYQPSVCSSTSNFPGAGVYSPQSHSYSGMGCPQHPFN- 1138
 Db 1086 PPLPLHEGPPRAPSGSYQPSVCSSTSNFPGAGVYSPQSHSYSGMGCPQHPFN- 1145

QY 1139 -----RGSGP 1143
 Db 1146 FFAYLYTPRYPGSGSQLISRPVKTSLPRDQNGLEVLIKEDAAELSSPTDSGSGP 1205
 QY 1144 APGRVYVLLNSLNDVAVCEKIKOTIEGLDOSMLPOYCTTIKANKINGVLAOCNIDELKKEM 1203
 Db 1206 APGRVYVLLNSLNDVAVCEKIKOTIEGLDOSMLPOYCTTIKANKINGVLAOCNIDELKKEM 1265
 QY 1204 NMNFGDWHLFRSTVLEKRNASHVPEPRFLSESSSGPAPRAGEPARARASHNELPTELS 1263
 Db 1266 NMNFGDWHLFRSTVLEKRNASHVPEPRFLSESSSGPAPRAGEPARARASHNELPTELS 1325
 QY 1264 SQTPYTLNFSPEELNTIGLDEGAPRHSNLSWQOTRTPLSLNSQDSSIETSKITDY 1323
 Db 1326 SQTPYTLNFSPEELNTIGLDEGAPRHSNLSWQOTRTPLSLNSQDSSIETSKITDY 1385
 QY 1324 QAEYRDAYREYIAQMSQLEGGPGSTTISGRSPHSTYVMQOSSGGSISNLEOEKGRS 1383
 Db 1386 QAEYRDAYREYIAQMSQLEGGPGSTTISGRSPHSTYVMQOSSGGSISNLEOEKGRS 1445
 QY 1384 EPRPDGRKSFMLKRGVDIDYSSSGVSTNDASPLDITEDEKSDQSGSKLLPGKKSER 1443
 Db 1446 EPRPDGRKSFMLKRGVDIDYSSSGVSTNDASPLDITEDEKSDQSGSKLLPGKKSER 1505
 QY 1444 SSLPOTDLTKLKGSLRQKLPSPDEDSGPEESONTPLADDKDKRAKGVYRVPKSPHS 1503
 Db 1506 SSLPOTDLTKLKGSLRQKLPSPDEDSGPEESONTPLADDKDKRAKGVYRVPKSPHS 1565
 QY 1504 AEPTRETIKAKAYLSDALDKKSDSGSVSSSSSNHSLHNEVADDSOLEKANLELBD 1563
 Db 1566 AEPTRETIKAKAYLSDALDKKSDSGSVSSSSSNHSLHNEVADDSOLEKANLELBD 1625
 QY 1564 DSHSGKRGIPHSISGLDPLIAMSICSEDKKSPSECSLIASSPEENMPACOKAYVLAFT 1623
 Db 1626 DSHSGKRGIPHSISGLDPLIAMSICSEDKKSPSECSLIASSPEENMPACOKAYVLAFT 1685
 QY 1624 PSTYTLNNSNAPNANONFDEMEGIRETSOVLTRSSSPNPTTIONENMLKSTHRSOR 1683
 Db 1686 PSTYTLNNSNAPNANONFDEMEGIRETSOVLTRSSSPNPTTIONENMLKSTHRSOR 1745
 QY 1684 SSTYRLSKDPELHAAASSESTGFGERESIL 1715
 Db 1746 SSTYRLSKDPELHAAASSESTGFGERESIL 1777

RESULT 2
 QY 09EOG6 PRELIMINARY; PRT; 1762 AA.
 AC 09EOG6;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-OCT-2001 (TREMblrel. 18, last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, last annotation update)
 DE KIDINS220.
 GN KIDINS220.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20568256; Pubmed=10998417;
 RA Iglesias T., Cabrera-Poch N., Mitchell M.P., Naven T.J., Rozenfurt E.,
 Schiavo G.:
 RT "Identification and cloning of Kidins220, a novel neuronal substrate
 of protein kinase D.".
 RL J. Biol. Chem. 275:40048-40056(2000).
 RP [2]
 RA Schiavo G.:
 RL Submitted (Feb-2000) to the EMBL/GenBank/DBJ databases.
 RP [3]
 RP SEQUENCE FROM N.A.

RA Mitchell M., Schiavo G.;
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF239045; AG35185.2;
 DR HSSP: P80144; 2MYO.
 DR InterPro: IPR002110; ANK.
 DR Pfam: PF00023; ank.12.
 DR PRINTS: PRO1415; ANKYRIN.
 DR SMART: SM00248; ANK.11.
 DR PROSITE: PS00088; ANK_REPEAT; 10.
 DR ANK repeat.
 KW ANK repeat.
 SO SEQUENCE 1762 AA; 195715 MW; 0CB2689A571F8A4 CRC64;

Query Match 94.1%; Score 8362.5; DB 11; Length 1762;
 Best Local Similarity 91.6%; Pred. No. 0;
 Matches 1616; Conservative 47; Mismatches 50; Indels 51; Gaps 4;

QY 1 MSVLSGSVINYEEENIPALKALECKYDNERNEGQTPMLAAEQGLYELKLN 60
 DB 1 MSVLSGSVINYEEENIPALKALECKYDNERNEGQTPMLAAEQGLYELKLN 60
 QY 61 GANCNLEDDLNMTALLISASKEGHVIEELKCGVNLHHRDMGWTALMAGYKGRDYY 120
 DB 61 GANCNLEDDLNMTALLISASKEGHVIEELKCGVNLHHRDMGWTALMAGYKGRDYY 120
 QY 121 ELLISHGANDSVTGLQYVYPIIWAAGRHADIVHLLONGARVNCSDXYTTPVWAAR 180
 DB 121 ELLISHGANDSVTGLQYVYPIIWAAGRHADIVHLLONGARVNCSDXYTTPVWAAR 180
 QY 181 KGHLECKHLLANGADVDOGANSMTALYAAGGYQSYKELKRPNNLTDKGNTA 240
 DB 181 KGHLECKHLLANGADVDOGANSMTALYAAGGYQSYKELKRPNNLTDKGNTA 240
 QY 180 KGHLECKHLLANGADVDOGANSMTALYAAGGYQSYKELKRPNNLTDKGNTA 239
 DB 241 LMIASKEGHEIVODLDAGTYNIPDRSGTYLIGAVRGHVEIYALLQKADIDIRG 300
 DB 240 LMIASKEGHEIVODLDAGTYNIPDRSGTYLIGAVRGHVEIYALLQKADIDIRG 299
 QY 301 QDNKTAALYMAVERKGNATVRDILQCNPDTEICTKDETPLIKATKMRNIEVEVELLDKGA 360
 DB 300 QDNKTAALYMAVERKGNATVRDILQCNPDTEICTKDETPLIKATKMRNIEVEVELLDKGA 359
 QY 361 KVSADVOKKGPPLHAIARGSRKIAELLNPNKDGRLLYPNKAGETPYNIDCSHOKSL 420
 DB 360 KVSADVOKKGPPLHAIARGSRKIAELLNPNKDGRLLYPNKAGETPYNIDCSHOKSL 419
 QY 421 TQIFGARHLSPTEEDGMLGYDLYSSALADILSEPTMQPICYGLYQWMSGKSFLLKTL 480
 DB 420 TQIFGARHLSPTEEDGMLGYDLYSSALADILSEPTMQPICYGLYQWMSGKSFLLKTL 479
 QY 481 EDEKRTFAGQOIEPLFOFSMLYVFLTLCCGGLGLLFAFTVHPNLGIAVSLFLALLYIF 540
 DB 480 EDEKRTFAGQOIEPLFOFSMLYVFLTLCCGGLGLLFAFTVHPNLGIAVSLFLALLYIF 539
 QY 541 FIYIFEGRRREGESWMAWVLSRFLARHIGYLELLKIMFVNPELPEQOTKALPVREF 600
 DB 540 FIYIFEGRRREGESWMAWVLSRFLARHIGYLELLKIMFVNPELPEQOTKALPVREF 599
 QY 601 TDVNRLLSSVGETSLAEMIALTLSDACEREGFLATRLFRVFKTEDTGKKMKWKTCLPS 660
 DB 600 TDVNRLLSSVGETSLAEMIALTLSDACEREGFLATRLFRVFKTEDTGKKMKWKTCLPS 659
 QY 661 FVIFLEFIIGCTISGTLTALFRVDPKHLTVNAVLSIASVGLAFVLCNRTWQVLDSTL 720
 DB 660 FVIFLEFIIGCTISGTLTALFRVDPKHLTVNAVLSIASVGLAFVLCNRTWQVLDSTL 719
 QY 721 NSQKRLHNAASKHAKLSGSEFMVYLKCEVELMARAKTIDSFQNTRLVVIIDGIDAC 780
 DB 720 NSQKRLHNAASKHAKLSGSEFMVYLKCEVELMARAKTIDSFQNTRLVVIIDGIDAC 779
 QY 781 EODVYLOMDTVRYLFEKGPFIATFASDPHIIITAIQONNSVLRDNINGHMYRIVH 840
 DB 780 EODVYLOMDTVRYLFEKGPFIATFASDPHIIITAIQONNSVLRDNINGHMYRIVH 839
 QY 841 LPVFLNSRGLSNARKFLVTSATNGDVPSCDPTGIGQEDADRRVSQNSLGEMTKLSKSTALN 900

DB 840 LPVFLNSRGLSNARKFLVTSATNGDVPSCDPTGIGQEDADRRVSQNSLGEMTKLSKSTALN 899
 QY 901 RBDYRRRMOORTIRROMSFOLTKLLYMEDMFSDSIPQTMRLNIVSVTGLLRANQIS 960
 DB 900 RBDYRRRMOORTIRROMSFOLTKLLYMEDMFSDSIPQTMRLNIVSVTGLLRANQIS 959
 QY 961 FVMDRLASWINTEQMPYTSWILYLETBESIPQOMTLKTYERISKNIPTTKVDEPL 1020
 DB 960 FVMDRLASWINTEQMPYTSWILYLETBESIPQOMTLKTYERISKNIPTTKVDEPL 1019
 QY 1021 EIDGDIRNEVEFLSRTPLVADYKVFPLCTVNLDPKIRELIADYRAAREQISIGLAY 1080
 DB 1020 EIDGDIRNEVEFLSRTPLVADYKVFPLCTVNLDPKIRELIADYRAAREQISIGLAY 1079
 QY 1081 PPLPLEGPPRPPSGSQSPVSSCSSTFNGPRAGVSVSPQSSSYSGSMGTQGHFYNR 1139
 DB 1080 PPLPLEGPPRPPSGSQSPVSSCSSTFNGPRAGVSVSPQSSSYSGSMGTQGHFYNR 1139
 QY 1140 -----GSG-----PAGPVVL 1151
 DB 1140 FPAPLYTTRYPPGSGQLHSRSVKTSLPQRONNGLPDDSGFNKROQAAPVATSSILL 1199
 QY 1152 NSLNDVAVCEKIQLEGDQSMPLPOYCTTIKKNANINGRYLAOCNIDELKKEKNNFQDW 1211
 DB 1200 SMTVDVYCEKIQLEGDQSMPLPOYCTTIKKNANINGRYLAOCNIDELKKEKNNFQDW 1259
 QY 1212 LFRSTYLBKRNASHVYPPDRPRLSSSSGAPRHPHAPARASHNLPHTELSSQTPYLN 1271
 DB 1260 LFRSMYLEKRSVSEQVPPDRPRLSSSSGAPRHPHAPARASHNLPHTELSSQTPYLN 1319
 QY 1272 FSPFEELNTGLDEGAPRHSNLSMSQOTRTPPLSSINSODSIEISKLTDKQAYERYAY 1331
 DB 1320 FSPFEELNTGLDEGAPRHSNLSMSQOTRTPPLSSINSODSIEISKLTDKQAYERYAY 1379
 QY 1332 REVIAQMSQLEGGPGSTTISGRSSPHSTYMYGQSSSGSISNLEQKGEKSEPRPDDGR 1391
 DB 1380 REVIAQMSQLEGGPGSTTISGRSSPHSTYMYGQSSSGSISNLEQKGEKSEPRPDDGR 1439
 QY 1392 KSTFLMRGVYIDYSSSGVSTNDASPLDPTTEDEKSDSGSKLLPGEKSSSEPSLFQIDL 1451
 DB 1440 KSTFLMRGVYIDYSSSGVSTNDASPLDPTTEDEKSDSGSKLLPGEKSSSEPSLFQIDL 1499
 QY 1452 KLGSGLRQKLPSPDEDESGTEESDNTPLKDKDKRKAAGYERVPKSEHAEPIRIFI 1511
 DB 1500 KLGSGLRQKLPSPDEDESGTEESDNTPLKDKDKRKAAGYERVPKSEHAEPIRIFI 1559
 QY 1512 KAKYVSDALLDKKSDSGSVRSSESSPNHSLHNEVADDSQLEKANLIELEDSSHGRG 1571
 DB 1560 KAKYVSDALLDKKSDSGSVRSSESSPNHSLHNEVADDSQLEKANLIELEDSSHGRG 1619
 QY 1572 IPHSLSGLODPIIARSTISEDKSSPESCSLIASSPEEWAPACOKAYNLNRPSTVTLLN 1631
 DB 1620 IPHSLSGLODPIIARSTISEDKSSPESCSLIASSPEEWAPACOKAYNLNRPSTVTLLN 1679
 QY 1632 NSAPANRANQNDDEMIGIRETSQVILRPSSSPPTTIQONENKSKTHRSQSSYTRLSK 1691
 DB 1680 NSAPANRANQNDDEMIGIRETSQVILRPSSSPPTTIQONENKSKTHRSQSSYTRLSK 1739
 QY 1692 DPPELHAAASSESTGGEERESIL 1715
 DB 1740 DASELH-AAASSESTGGEERESIL 1762

RESULT 3
 Q9ERD4 PRELIMINARY; PRT: 1715 AA.
 AC Q9ERD4;
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Ankyrin repeat-rich membrane-spanning protein.
 GN ARMS.

OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euteria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OC NCBI_TaxID=10116;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY;
 RX MEDLINE=20585245; PubMed=1150334;
 RA Kong H., Boulter J., Weber J.L., Lai C., Chao M.V.;
 RT "An evolutionarily conserved transmembrane protein that is a novel
 downstream target of neurotrophin and ephrin receptors."
 RL J. Neurosci. 21:176-185(2001).
 DR EMBL: AF313464; AAG34167.1; .
 DR HSSP: P80144; 2MYO.
 DR InterPro: IPR002110; ANK.
 DR Pfam: PF00023; ank; 12.
 DR PRINTS: PRO1415; ANKTRIN.
 DR SMART: SM00248; ANK; 11.
 DR PROSITE: PS50088; ANK_REPEAT; 10.
 DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
 DR ANK repeat. Repeat.
 DR SEQUENCE 1715 AA; 190532 MW; 505593C4A19A6CDC CRC64;

Query Match 93.0%; Score 8266; DB 11; Length 1715;
 Best Local Similarity 92.7%; Pred. No. 0;
 Matches 1591; Conservative 53; Mismatches 70; Indels 2; Gaps 2;

QY 1 MSVLISQSYINVEENIPALKEKCKDVERNEGOTPLMTAAEGNLEIKELKN 60
 DB 1 MSVLISQSYINVEENIPALKEKCKDVERNEGOTPLMTAAEGNLEIKELKN 60
 QY 61 GANCMIEDLDNMTALISAKSGHVIIVELLKCGVNEHRMGWNTALMAACGRDGV 120
 DB 61 GANCMIEDLDNMTALISAKSGHVIIVELLKCGVNEHRMGWNTALMAACGRDGV 120
 QY 121 ELLSHGANPSTYTGQYVYPIYMAAGRHADYHLLQNGAKVNSKRTGTPLYMAAR 180
 DB 121 ELLSHGANPSTYTGQYVYPIYMAAGRHADYHLLQNGAKVNSKRTGTPLYMAAR 180
 QY 121 ELLSHGANPSTYTGQYVYPIYMAAGRHADYHLLQNGAKVNSKRTGTPLYMAAR 180
 DB 121 ELLSHGANPSTYTGQYVYPIYMAAGRHADYHLLQNGAKVNSKRTGTPLYMAAR 180
 QY 181 KGHLECVKHLAMGADVQOEGANSMTALIVAKGQYTSVKEILKRNVMJTRKDN 240
 DB 181 KGHLECVKHLAMGADVQOEGANSMTALIVAKGQYTSVKEILKRNVMJTRKDN 240
 QY 181 KGHLECVKHLAMGADVQOEGANSMTALIVAKGQYTSVKEILKRNVMJTRKDN 240
 DB 181 KGHLECVKHLAMGADVQOEGANSMTALIVAKGQYTSVKEILKRNVMJTRKDN 240
 QY 241 LMIASKEGHTETVDDLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLOK 300
 DB 241 LMIASKEGHTETVDDLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLOK 300
 QY 241 LMIASKEGHTETVDDLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLOK 300
 DB 241 LMIASKEGHTETVDDLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLOK 300
 QY 301 QDNKTAALYMAVEKGNATVVRDILQCNPTETCTDGETPLKATKRMRIEYVELLD 360
 DB 301 QDNKTAALYMAVEKGNATVVRDILQCNPTETCTDGETPLKATKRMRIEYVELLD 360
 QY 301 QDNKTAALYMAVEKGNATVVRDILQCNPTETCTDGETPLKATKRMRIEYVELLD 360
 DB 301 QDNKTAALYMAVEKGNATVVRDILQCNPTETCTDGETPLKATKRMRIEYVELLD 360
 QY 361 KVSANDKKGDTPLAIAIGRSRKLAEILLRNPDKDRLLYRNKAGETPTNIDCSH 420
 DB 361 KVSANDKKGDTPLAIAIGRSRKLAEILLRNPDKDRLLYRNKAGETPTNIDCSH 420
 QY 361 KVSANDKKGDTPLAIAIGRSRKLAEILLRNPDKDRLLYRNKAGETPTNIDCSH 420
 DB 361 KVSANDKKGDTPLAIAIGRSRKLAEILLRNPDKDRLLYRNKAGETPTNIDCSH 420
 QY 421 TQIFGARHLSPETDGMGYDLYSSALADILSEPTMOPICVGLYAQMGSGKSFLL 480
 DB 421 TQIFGARHLSPETDGMGYDLYSSALADILSEPTMOPICVGLYAQMGSGKSFLL 480
 QY 421 TQIFGARHLSPETDGMGYDLYSSALADILSEPTMOPICVGLYAQMGSGKSFLL 480
 DB 421 TQIFGARHLSPETDGMGYDLYSSALADILSEPTMOPICVGLYAQMGSGKSFLL 480
 QY 481 EDEMTFAGQOETPLFQFSWLVLTLLCGGLGFVFAFVADTNALAISSFLALYIF 540
 DB 481 EDEMTFAGQOETPLFQFSWLVLTLLCGGLGFVFAFVADTNALAISSFLALYIF 540
 QY 481 EDEMTFAGQOETPLFQFSWLVLTLLCGGLGFVFAFVADTNALAISSFLALYIF 540
 DB 481 EDEMTFAGQOETPLFQFSWLVLTLLCGGLGFVFAFVADTNALAISSFLALYIF 540
 QY 541 FIVIFGGRGEGSNMAMVLSLRLAIGLYELLKLMFNPELPBOTTKALPVPLF 600
 DB 541 FIVIFGGRGEGSNMAMVLSLRLAIGLYELLKLMFNPELPBOTTKALPVPLF 600
 QY 541 FIVIFGGRGEGSNMAMVLSLRLAIGLYELLKLMFNPELPBOTTKALPVPLF 600
 DB 541 FIVIFGGRGEGSNMAMVLSLRLAIGLYELLKLMFNPELPBOTTKALPVPLF 600
 QY 601 TDYNLSSVGGSTLAEMIATLSACEREFGLATRLRVKTEDTQGGKKWKTCCPLS 660
 DB 601 TDYNLSSVGGSTLAEMIATLSACEREFGLATRLRVKTEDTQGGKKWKTCCPLS 660
 QY 601 TDYNLSSVGGSTLAEMIATLSACEREFGLATRLRVKTEDTQGGKKWKTCCPLS 660
 DB 601 TDYNLSSVGGSTLAEMIATLSACEREFGLATRLRVKTEDTQGGKKWKTCCPLS 660
 QY 661 FVIFLFIIGCIISGTTLLAIFRVPKHLTVNAVILSISVGLAFVLCNCRWMOYDLSL 720
 DB 661 FVIFLFIIGCIISGTTLLAIFRVPKHLTVNAVILSISVGLAFVLCNCRWMOYDLSL 720

DB 661 FVIFLFIIGCIISGTTLLAIFRVPKHLTVNAVILSISVGLAFVLCNCRWMOYDLSL 720
 QY 721 NSQKRLHNAASLHLKIKSEGMKVLKCEYELMARAKTIIDSTONQOTRLVYIDGLAC 780
 DB 721 NSQKRLHNAASLHLKIKSEGMKVLKCEYELMARAKTIIDSTONQOTRLVYIDGLAC 780
 QY 721 NSQKRLHNAASLHLKIKSEGMKVLKCEYELMARAKTIIDSTONQOTRLVYIDGLAC 780
 DB 721 NSQKRLHNAASLHLKIKSEGMKVLKCEYELMARAKTIIDSTONQOTRLVYIDGLAC 780
 QY 781 EODKVLQMDTVAVLSKGFIFAFSDPIIITKALNOUNLSVLRSGNNGHYMRNYH 840
 DB 781 EODKVLQMDTVAVLSKGFIFAFSDPIIITKALNOUNLSVLRSGNNGHYMRNYH 840
 QY 781 EODKVLQMDTVAVLSKGFIFAFSDPIIITKALNOUNLSVLRSGNNGHYMRNYH 840
 DB 781 EODKVLQMDTVAVLSKGFIFAFSDPIIITKALNOUNLSVLRSGNNGHYMRNYH 840
 QY 841 LPVFLNSRGLSNARKFLVTSATNGDPCSDTGTIOBADRVSONSLGEMTKSGTALN 900
 DB 841 LPVFLNSRGLSNARKFLVTSATNGDPCSDTGTIOBADRVSONSLGEMTKSGTALN 900
 QY 841 LPVFLNSRGLSNARKFLVTSATNGDPCSDTGTIOBADRVSONSLGEMTKSGTALN 900
 DB 841 LPVFLNSRGLSNARKFLVTSATNGDPCSDTGTIOBADRVSONSLGEMTKSGTALN 900
 QY 901 RRDYRRRQKRTITQMSFDLKLVTEDWEDISDIPQMRRLNIVSTGRLRANQIS 960
 DB 901 RRDYRRRQKRTITQMSFDLKLVTEDWEDISDIPQMRRLNIVSTGRLRANQIS 960
 QY 901 RRDYRRRQKRTITQMSFDLKLVTEDWEDISDIPQMRRLNIVSTGRLRANQIS 960
 DB 901 RRDYRRRQKRTITQMSFDLKLVTEDWEDISDIPQMRRLNIVSTGRLRANQIS 960
 QY 961 FNMDRLASWINTLEQMPYRTSWLILYLEETEGIPDQMTLKTIERISKNIPTTKDVEPL 1020
 DB 961 FNMDRLASWINTLEQMPYRTSWLILYLEETEGIPDQMTLKTIERISKNIPTTKDVEPL 1020
 QY 961 FNMDRLASWINTLEQMPYRTSWLILYLEETEGIPDQMTLKTIERISKNIPTTKDVEPL 1020
 DB 961 FNMDRLASWINTLEQMPYRTSWLILYLEETEGIPDQMTLKTIERISKNIPTTKDVEPL 1020
 QY 1021 EIDGDIRNEFVLSSRTPVLVARDVFLPCTVNLDPKUREITADYRAAREQISIGCLAY 1080
 DB 1021 EIDGDIRNEFVLSSRTPVLVARDVFLPCTVNLDPKUREITADYRAAREQISIGCLAY 1080
 QY 1021 EIDGDIRNEFVLSSRTPVLVARDVFLPCTVNLDPKUREITADYRAAREQISIGCLAY 1080
 DB 1021 EIDGDIRNEFVLSSRTPVLVARDVFLPCTVNLDPKUREITADYRAAREQISIGCLAY 1080
 QY 1081 PPLPIHGRPPRPSGYSQPPSVCSMTSPNPFPGVAVSQPHSSYISGSLSGDHPYRNA 1140
 DB 1081 PPLPIHGRPPRPSGYSQPPSVCSMTSPNPFPGVAVSQPHSSYISGSLSGDHPYRNA 1140
 QY 1081 PPLPIHGRPPRPSGYSQPPSVCSMTSPNPFPGVAVSQPHSSYISGSLSGDHPYRNA 1140
 DB 1081 PPLPIHGRPPRPSGYSQPPSVCSMTSPNPFPGVAVSQPHSSYISGSLSGDHPYRNA 1140
 QY 1141 SGAPGPVYVLSLNDVAVCEKLOJEGIDQSMIPQYCTTIKANINGVLAQCNIDELK 1200
 DB 1141 SGAPGPVYVLSLNDVAVCEKLOJEGIDQSMIPQYCTTIKANINGVLAQCNIDELK 1200
 QY 1141 SGAPGPVYVLSLNDVAVCEKLOJEGIDQSMIPQYCTTIKANINGVLAQCNIDELK 1200
 DB 1141 SGAPGPVYVLSLNDVAVCEKLOJEGIDQSMIPQYCTTIKANINGVLAQCNIDELK 1200
 QY 1201 KEMAMNGDHLFRSYVLEKRNASHVVPDPRFLESSESGAPRHPGEPARRASHNELPPT 1260
 DB 1201 KEMAMNGDHLFRSYVLEKRNASHVVPDPRFLESSESGAPRHPGEPARRASHNELPPT 1260
 QY 1201 KEMAMNGDHLFRSYVLEKRNASHVVPDPRFLESSESGAPRHPGEPARRASHNELPPT 1260
 DB 1201 KEMAMNGDHLFRSYVLEKRNASHVVPDPRFLESSESGAPRHPGEPARRASHNELPPT 1260
 QY 1261 ELSQTPPYTLNFSFEELNTGLDEGAPRHSNLSWQOTRRTPSLSLNSQDSIEISKLT 1320
 DB 1261 ELSQTPPYTLNFSFEELNTGLDEGAPRHSNLSWQOTRRTPSLSLNSQDSIEISKLT 1320
 QY 1261 ELSQTPPYTLNFSFEELNTGLDEGAPRHSNLSWQOTRRTPSLSLNSQDSIEISKLT 1320
 DB 1261 ELSQTPPYTLNFSFEELNTGLDEGAPRHSNLSWQOTRRTPSLSLNSQDSIEISKLT 1320
 QY 1321 DKVQAEYRDYREYIAQMSOLEGPGSTTISGRSSPHSTYYMGQSSGGSIHNSLEQEK 1380
 DB 1321 DKVQAEYRDYREYIAQMSOLEGPGSTTISGRSSPHSTYYMGQSSGGSIHNSLEQEK 1380
 QY 1321 DKVQAEYRDYREYIAQMSOLEGPGSTTISGRSSPHSTYYMGQSSGGSIHNSLEQEK 1380
 DB 1321 DKVQAEYRDYREYIAQMSOLEGPGSTTISGRSSPHSTYYMGQSSGGSIHNSLEQEK 1380
 QY 1381 KDSEPKPDGGRKSEFLKRGVYIDYSSGVSTNDASPLDITTEDEKSDQSGSKILPGK 1440
 DB 1381 KDSEPKPDGGRKSEFLKRGVYIDYSSGVSTNDASPLDITTEDEKSDQSGSKILPGK 1440
 QY 1381 KDSEPKPDGGRKSEFLKRGVYIDYSSGVSTNDASPLDITTEDEKSDQSGSKILPGK 1440
 DB 1381 KDSEPKPDGGRKSEFLKRGVYIDYSSGVSTNDASPLDITTEDEKSDQSGSKILPGK 1440
 QY 1441 SERSSLEFQDYLKLGSGLRQKLPDEDSGTEESNPT-LTKDKRAAEKVERVPS 1499
 DB 1441 SERSSLEFQDYLKLGSGLRQKLPDEDSGTEESNPT-LTKDKRAAEKVERVPS 1499
 QY 1441 SERSSLEFQDYLKLGSGLRQKLPDEDSGTEESNPT-LTKDKRAAEKVERVPS 1499
 DB 1441 SERSSLEFQDYLKLGSGLRQKLPDEDSGTEESNPT-LTKDKRAAEKVERVPS 1499
 QY 1500 PERSAEPRTITAKAYLSDALLDKRSDSGVRSSESPNHSANEVADDSOLEKANI 1559
 DB 1500 PERSAEPRTITAKAYLSDALLDKRSDSGVRSSESPNHSANEVADDSOLEKANI 1559
 QY 1500 PERSAEPRTITAKAYLSDALLDKRSDSGVRSSESPNHSANEVADDSOLEKANI 1559
 DB 1500 PERSAEPRTITAKAYLSDALLDKRSDSGVRSSESPNHSANEVADDSOLEKANI 1559
 QY 1501 QEHSABEIRTFIAKAYLSDALLDKRSDSGVRSSESPNHSANEVADDSOLEKANI 1560
 DB 1501 QEHSABEIRTFIAKAYLSDALLDKRSDSGVRSSESPNHSANEVADDSOLEKANI 1560
 QY 1560 ELEDHSHSGRGIPIHLSIGLOPITARMGICSDKSPESCSLIASSPEENPACOKAN 1619
 DB 1560 ELEDHSHSGRGIPIHLSIGLOPITARMGICSDKSPESCSLIASSPEENPACOKAN 1619
 QY 1561 ELEDHSHSGRGIPIHLSIGLOPITARMGICSDKSPESCSLIASSPEENPACOKAN 1620
 DB 1561 ELEDHSHSGRGIPIHLSIGLOPITARMGICSDKSPESCSLIASSPEENPACOKAN 1620
 QY 1620 LNTPTSTVTLNNSASANRANQNFDMEGIRETSQYLTPSSSPNPTTQONELSMTHK 1679
 DB 1620 LNTPTSTVTLNNSASANRANQNFDMEGIRETSQYLTPSSSPNPTTQONELSMTHK 1679
 QY 1621 LNTPTSTVTLNNTATNTANQNFDEIGIRETSQYLTPGSPNPTTAQONELSMTHK 1680
 DB 1621 LNTPTSTVTLNNTATNTANQNFDEIGIRETSQYLTPGSPNPTTAQONELSMTHK 1680
 QY 1680 RSQSSYTRLSKDPPELAAAASESTGFBERSIL 1715
 DB 1680 RSQSSYTRLSKDPPELAAAASESTGFBERSIL 1715
 QY 1681 RSQSSYTRLSKDPPELAAAASESTGFBERSIL 1715
 DB 1681 RSQSSYTRLSKDPPELAAAASESTGFBERSIL 1715

RESULT 4
 Q90F42

ID 09UF42 PRELIMINARY; PRT; 1031 AA.
 AC 09UF42;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Hypothetical 115.3 kDa protein.
 GN DKEP344P0621.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Tissue-tests;
 RL Blum H., Baerends S., Mewes H.W., Gassenhuber J., Wiemann S.;
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL133620; CAB63746.1;
 DR HSSP; P80144; 2MYO.
 DR InterPro; IPR002110; ANK.
 DR Pfam; PF00023; ank. 12.
 DR PRINTS; PR01415; ANKRN.
 DR SMART; SM00248; ANK. 11.
 DR PROSITE; PS50088; ANK_REPEAT. 10.
 DR PROSITE; PS50297; ANK_REPEAT_REGION. 1.
 DR ANK repeat: Hypothetical protein; Repeat.
 SQ SEQUENCE 1031 AA; 115311 MW; D9E866461C13A2F5 CRC64;

Query Match 58.6%; Score 5205.5; DB 4; Length 1031;
 Best Local Similarity 98.5%; Pred. No. 0;
 Matches 1008; Conservative 2; Mismatches 0; Indels 13; Gaps 1;

QY 1 MSVLSQSVINVEEENIPKALKECKDVDERNECGOTPLIAAEOGULEYVKEIKN 60
 DB 1 MSVLSQSVINVEEENIPKALKECKDVDERNECGOTPLIAAEOGULEYVKEIKN 60
 QY 61 GANCNLEDDLNMTALISASKEGVHIVEELKCGVNLHRDMGWTALMAGYKGRDYY 120
 DB 61 GANCNLEDDLNMTALISASKEGVHIVEELKCGVNLHRDMGWTALMAGYKGRDYY 120
 QY 121 ELLSHGANSVYGLQSVYPIIWAARGHADIVHLLQNGAVNCSDKGTPIYMAAR 180
 DB 121 ELLSHGANSVYGLQSVYPIIWAARGHADIVHLLQNGAVNCSDKGTPIYMAAR 180
 QY 181 KGHLECYKHLIANGADVDGANSMTALIAVAGGYQSYKEILKRNPNVLTDKDQNTA 240
 DB 181 KGHLECYKHLIANGADVDGANSMTALIAVAGGYQSYKEILKRNPNVLTDKDQNTA 240
 QY 241 LMIASKEGHEIYODLDAGTYVNIPIRSGDTVLIGAVRGHVEIYRALLQKYADIDIRG 300
 DB 241 LMIASKEGHEIYODLDAGTYVNIPIRSGDTVLIGAVRGHVEIYRALLQKYADIDIRG 300
 QY 301 QDNKNTALYMAVEKGNATVNDIIQCNPDTEICRKDGTPILKATKMNIEVVELLDKGA 360
 DB 301 QDNKNTALYMAVEKGNATVNDIIQCNPDTEICRKDGTPILKATKMNIEVVELLDKGA 360
 QY 361 KVSADVKKDPTPLHAIIRGRSRKLAELLRNPKDRLLYPNKAGETPNYNDSSHOKSIL 420
 DB 361 KVSADVKKDPTPLHAIIRGRSRKLAELLRNPKDRLLYPNKAGETPNYNDSSHOKSIL 420
 QY 421 TQIFGARHLSPTEDGMGLDYLSALADILSEPTQPIQVGLVAQWGSGLKLLKL 480
 DB 421 TQIFGARHLSPTEDGMGLDYLSALADILSEPTQPIQVGLVAQWGSGLKLLKL 480
 QY 481 EDEKTRAGOOIEPLFQFSMLIVFLTLGGLGLFAFTVHPMLGIAVSLFALLIYF 540
 DB 481 EDEKTRAGOOIEPLFQFSMLIVFLTLGGLGLFAFTVHPMLGIAVSLFALLIYF 540
 QY 541 FIVYFGRRGREGESMNAWVLTSLARHIGYLELLKLMFVNPELPEQOTTKALPVLELF 600
 DB 541 FIVYFGRRGREGESMNAWVLTSLARHIGYLELLKLMFVNPELPEQOTTKALPVLELF 600
 QY 601 TDYARLSSVGGESTLAMIATLSDACEREGFLATRLFRVFKPDNDGKKKKKTCCLPS 660
 DB 601 TDYARLSSVGGESTLAMIATLSDACEREGFLATRLFRVFKPDNDGKKKKKTCCLPS 660

DB 601 TDYARLSSVGGESTLAMIATLSDACEREGFLATRLFRVFKPDNDGKKKKKTCCLPS 660
 QY 661 FVIEFLITGCTISGTTLAIFRVPKHLTVANAIVISTASVGLAFVLCRTMGOVLSL 720
 DB 661 FVIEFLITGCTISGTTLAIFRVPKHLTVANAIVISTASVGLAFVLCRTMGOVLSL 720
 QY 721 NSQRRRLNNAASKLKLKSEGFPMKLVCEVELMARMATIDISFTQNTRLVVIDGLDAC 780
 DB 721 NSQRRRLNNAASKLKLKSEGFPMKLVCEVELMARMATIDISFTQNTRLVVIDGLDAC 780
 QY 781 EODKVLQMLDYRVVLFSGKPTAIFASDPHIIKAINONLSVLRDSNSHDMYRNIVH 840
 DB 781 EODKVLQMLDYRVVLFSGKPTAIFASDPHIIKAINONLSVLRDSNSHDMYRNIVH 840
 QY 841 LPVFLNSGLSARKEIVTSATNGVPCSDTGTIOEDDADRVRSQNSLGEMLKLSKALN 900
 DB 841 LPVFLNSGLSARKEIVTSATNGVPCSDTGTIOEDDADRVRSQNSLGEMLKLSKALN 900
 QY 901 RRDYRRRQMQRTITRQMSFDLTKLVTEDWFSDISPQTMRLNIVSVTGRLLRANQIS 960
 DB 901 RRDYRRRQMQRTITRQMSFDLTKLVTEDWFSDISPQTMRLNIVSVTGRLLRANQIS 960
 QY 961 FWMDRASWINITEQMPYRTSWLLIYLETBISIPQMTLKTIYER-----IS 1007
 DB 961 FWMDRASWINITEQMPYRTSWLLIYLETBISIPQMTLKTIYER-----IS 1007
 QY 1008 KNI 1010
 DB 1021 KSV 1023

RESULT 5

Q9H9E4 PRELIMINARY; PRT; 543 AA.

ID 09H9E4;
 AC 09H9E4;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
 DE CDNA FLJ12811 f1s, clone NT2RP2002475.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Isogai T., Ota T., Hayaashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
 RA Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,
 RA Nakamura Y., Nagahara K., Masuno Y., Sasaki N.,
 RA "NEDO human cDNA sequencing project."
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK022873; BAB14285.1;
 SQ SEQUENCE 543 AA; 59976 MW; F5D643D5A20C641D CRC64;

Query Match 31.4%; Score 2791; DB 4; Length 543;
 Best Local Similarity 99.6%; Pred. No. 7; 6e-169;
 Matches 541; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1173 MLPOCYTTIKKANINGRVLACQNDLKKEMMNFGDMILFRSTVLEMRNASHVYVPEP 1232
 DB 1 MLPOCYTTIKKANINGRVLACQNDLKKEMMNFGDMILFRSTVLEMRNASHVYVPEP 1232
 QY 1233 RFLSESSSGPAPGEPARRASHNELPHTELSSQPTVTLNFSPEELNTGLDGAHRHNL 1292
 DB 1233 RFLSESSSGPAPGEPARRASHNELPHTELSSQPTVTLNFSPEELNTGLDGAHRHNL 1292
 QY 1293 SWSQTRRRPSSLSLSQSSSIFISKLTQKVAEYRDARREIYLAQMSQLEGPGSTTISG 1352
 DB 1293 SWSQTRRRPSSLSLSQSSSIFISKLTQKVAEYRDARREIYLAQMSQLEGPGSTTISG 1352
 QY 1353 RSPSHSTYMGSSSGSISHSNLEQKDSPPKPDGKRSPLMRKGVYIDVSSSGVSTN 1412
 DB 1353 RSPSHSTYMGSSSGSISHSNLEQKDSPPKPDGKRSPLMRKGVYIDVSSSGVSTN 1412
 QY 1412 RSPSHSTYMGSSSGSISHSNLEQKDSPPKPDGKRSPLMRKGVYIDVSSSGVSTN 1412
 DB 1412 RSPSHSTYMGSSSGSISHSNLEQKDSPPKPDGKRSPLMRKGVYIDVSSSGVSTN 1412

```

OY 1413 DASPLDPTDEDEKSDGSKLLPKGKSSRSLSFQNDLTKGSGLRQKLPSPDESGT 1472
DB 241 DASPLDPTDEDEKSDGSKLLPKGKSSRSLSFQNDLTKGSGLRQKLPSPDESGT 300
OY 1473 ESDNTPLDKDKRAKAGKRYVVKSPHSAEPIRFIRAKEXYLDALDKKSPSGV 1532
DB 301 ESDNTPLDKDKRAKAGKRYVVKSPHSAEPIRFIRAKEXYLDALDKKSPSGV 360
OY 1533 RSSSESPHSHLNEVADDSOLEKANLLEEDSHSGKGIPIHSLSGLDPIIAMSICSE 1592
DB 361 RSSSESPHSHLNEVADDSOLEKANLLEEDSHSGKGIPIHSLSGLDPIIAMSICSE 420
OY 1593 DKSSSECSLIASSPEENWPAKOKAYNLRPTPTVTLNNNSAPNANONDEMEGIRET 1652
DB 421 DKSSSECSLIASSPEENWPAKOKAYNLRPTPTVTLNNNSAPNANONDEMEGIRET 480
OY 1653 SOVILRPSSSPNPTTIQENLKSMTHKRSQSSYTRLKSKDPELHAAASSTGFGERE 1712
DB 481 SOVILRPSSSPNPTTIQENLKSMTHKRSQSSYTRLKSKDPELHAAASSTGFGERE 540
DB 1713 STL 1715
DB 541 STL 543

```

RESULT 6
09W211 PRELIMINARY: PRT: 1498 AA.

AC 09W211 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
GN CG10074 protein.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
RX NCBI_TaxID=7227;
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE=20196006; PubMed=107311132;
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasly E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Botkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos R., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup I.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Dubin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howard T.J., Wei M.-H., Idegawa C.,
RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Mekulov G., Milshina N.V., Mobarry B., Morris J., Mosnelli A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheefer F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

```

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Styrbas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
RT "The genome sequence of Drosophila melanogaster";  
RL Science 287:2185-2195(2000).  
DR EMBL: AE003453; AAF46710.1; -.  
DR HSSP: P42773; 11HB.  
DR FLYBASE: FBgn0034615; CG10074.  
DR InterPro: IPR002110; ANK.  
DR InterPro: IPR000847; HTH_LysR.  
DR Pfam: PF00023; ANK_11.  
DR SMART: SM00248; ANK; 10.  
DR PROSITE: PS50088; ANK_REPEAT; 9.  
DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.  
DR PROSITE: PS00044; HTH_LysR_FAMILY; UNKNOWN_1.  
KW ANK repeat; Repeat.  
SQ SEQUENCE 1498 AA; 165031 MW; 7C2A888E1B64982C CRC64;  
Query Match 26.8%; Score 2379.5; DB 5; Length 1498;  
Best Local Similarity 37.6%; Pred. No. 6,1e-142;  
Matches 559; Conservative 259; Mismatches 490; Indels 179; Gaps 39;  
OY 1 MSYLISQSYINVEENIPALKALE-KKDYDERNECGQPLMTAEDGNEIYKELIK 59  
DB 23 MGLGRLRALLOYIDNNDISGLRALIDSRLLTDIDRENAVTVLAVVAGGAFRAFEFLA 82  
OY 60 NGANCLMEDLDMMTALISAKSGHVIYELIKCGNLEHRMGMTALMAACYGRTDY 119  
DB 83 RGDVQAEEDLDMMTALISAKSGHGHLDVQGLLDHGAEPHRRMGMTSLMAAYGHEEL 142  
OY 120 VELLISGANPSVTGLQSYVPIIWAAGRHADIVHLLQNGAKVNSDKGTPELVAA 179  
DB 143 VRLLDGKADGANHG-NYHIGALLMAAGRGYDIYELLVQRCAGKAVNVDKGTALVWAC 201  
OY 180 RKGHECVKHLAMGADVQEGANSMTALIVAKGQSYKELIKRPNVNLDPKDNF 239  
DB 202 RRGNEIVPTLLKAGANVTAGMTSWTPLVAAAGGHTDGVSSILLEKRPVNAADLKDNT 261  
OY 240 ALMIASKEGTEIVDLDAGTYVNIPIKSGDTVLIGAVRGHVEIVRALLQKYADIDR 299  
DB 262 ALCTASREGFODIASLIAAGAVINIDRGATPLIHVAKGRVVEALLKKNADVDIO 321  
OY 300 GQDNKTALWAVEKGNATVNRDILQNDPTEICTYDGETPLIKATKRNIEVELLDKG 359  
DB 322 GDRKTAIYTAVEKGTTPVKKLLATNPDESATKDTPLILAVARNRNLLEIVHLLDRK 381  
OY 360 AVSAVDKGDPTPHIATIGRSKLAELLNRPKGRLLYRNKGETPVNIDCSHOKSI 419  
DB 382 AVTASDKRGDTCILHIAARAKSTIVLEALLNPKKSQLLIRANKAGETPVNIDSHQTTI 441  
OY 420 LQVIFGARHLSTETEDGMGLDYLSALADLISPTQMPPICVGLAQMGSGKSFLLK 479  
DB 442 LQVIFGARHLSTETEDGMGLDYLSALADLISPTQMPPICVGLAQMGSGKSFLLK 501  
OY 480 LEDDKTPAGQOIEPLPQSMILVLTLLGCGGLGLAFATYPMUGLAVSISF 533  
DB 502 LEDDKTPAGQOIEPLPQSMILVLTLLGCGGLGLAFATYPMUGLAVSISF 556  
OY 534 LALVYIFVYIFGGRGEGSNMAMVLSLRAHIGVLELLKLMFVNPELDEQTTKA 593  
DB 557 LALVYIFVYIFGGRGEGSNMAMVLSLRAHIGVLELLKLMFVNPELDEQTTKA 613  
OY 594 LPVRLFTDYNRLSSVGETSLAEMIAVLSDACEREFGLATRLFRVETKE-DTQKK 651  
DB 614 KVRHFHFAEANSASPT-GDGAVALHMLALDAIESHYGLATRLYRAFRPCKLVGDWR 672  
OY 652 WKKTCCLSFYVFLFICIGISGTLIAIF-----RDPKHLTYNAVLISASVGLAFV 706  
DB 673 WKKTCCLSFYVFLFICIGISGTLIAIF-----RDPKHLTYNAVLISASVGLAFV 730  
673 WRRMCCIPVILFELALVTAVTGISLTVAVTFADKEKEHILV-ALYVIAAVAGTLC 730

```


ID	099407;	PRELIMINARY;	PRT; 1856 AA.
AC	099407;		
DT	01-MAY-1997 (TREMBLrel. 03, Created)		
DT	01-MAY-1997 (TREMBLrel. 03, Last sequence update)		
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)		
DE	Ankyrin.		
GN	ANK1.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
OX	NCBI_taxid=9606;		
RN	{1}		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=97382244; Pubmed=9235914;		
RA	Callagher P.G., Tse W.T., Scarpa A.L., Lux S.E., Forget B.G.;		
RT	"Structure and organization of the human ankyrin-1 gene. Basis for		
RL	J. Biol. Chem. 272:19220-19228(1997).		
DR	EMBL; U50133; AAB47805.1; -		
DR	EMBL; U50092; AAB47805.1; JOINED.		
DR	EMBL; U50093; AAB47805.1; JOINED.		
DR	EMBL; U50095; AAB47805.1; JOINED.		
DR	EMBL; U50096; AAB47805.1; JOINED.		
DR	EMBL; U50097; AAB47805.1; JOINED.		
DR	EMBL; U50098; AAB47805.1; JOINED.		
DR	EMBL; U50099; AAB47805.1; JOINED.		
DR	EMBL; U50100; AAB47805.1; JOINED.		
DR	EMBL; U50101; AAB47805.1; JOINED.		
DR	EMBL; U50102; AAB47805.1; JOINED.		
DR	EMBL; U50103; AAB47805.1; JOINED.		
DR	EMBL; U50104; AAB47805.1; JOINED.		
DR	EMBL; U50105; AAB47805.1; JOINED.		
DR	EMBL; U50106; AAB47805.1; JOINED.		
DR	EMBL; U50107; AAB47805.1; JOINED.		
DR	EMBL; U50108; AAB47805.1; JOINED.		
DR	EMBL; U50109; AAB47805.1; JOINED.		
DR	EMBL; U50110; AAB47805.1; JOINED.		
DR	EMBL; U50111; AAB47805.1; JOINED.		
DR	EMBL; U50112; AAB47805.1; JOINED.		
DR	EMBL; U50113; AAB47805.1; JOINED.		
DR	EMBL; U50114; AAB47805.1; JOINED.		
DR	EMBL; U50115; AAB47805.1; JOINED.		
DR	EMBL; U50116; AAB47805.1; JOINED.		
DR	EMBL; U50117; AAB47805.1; JOINED.		
DR	EMBL; U50118; AAB47805.1; JOINED.		
DR	EMBL; U50119; AAB47805.1; JOINED.		
DR	EMBL; U50120; AAB47805.1; JOINED.		
DR	EMBL; U50121; AAB47805.1; JOINED.		
DR	EMBL; U50122; AAB47805.1; JOINED.		
DR	EMBL; U50123; AAB47805.1; JOINED.		
DR	EMBL; U50124; AAB47805.1; JOINED.		
DR	EMBL; U50125; AAB47805.1; JOINED.		
DR	EMBL; U50126; AAB47805.1; JOINED.		
DR	EMBL; U50127; AAB47805.1; JOINED.		
DR	EMBL; U50128; AAB47805.1; JOINED.		
DR	EMBL; U50129; AAB47805.1; JOINED.		
DR	EMBL; U50130; AAB47805.1; JOINED.		
DR	EMBL; U50131; AAB47805.1; JOINED.		
DR	EMBL; U50132; AAB47805.1; JOINED.		
DR	HSSP; O00420; IAWC.		
DR	InterPro: IPR002110; ANK.		
DR	InterPro: IPR000488; Death.		
DR	InterPro: IPR000306; ZUS.		
DR	Pfam: PF00023; ank; 23.		
DR	Pfam: PF00531; death; 1.		
DR	Pfam: PF00791; ZUS; 1.		
DR	PRINTS: PR01415; ANKYRIN.		
DR	SMART: SM00248; ANK; 21.		
DR	SMART: SM00005; DEATH; 1.		
DR	SMART; SM00218; ZUS; 1.		
DR	PROSITE: PS50088; ANK_REPEAT; 20.		
DR	PROSITE: PS50297; ANK_REP_REGION; 1.		

[illegible]

```

Db      913 IPRPCAPTRITCRFLVQPKLSTPPPLAEFEGSLASRIALGPTAQFLSTPIVTEIPHA 972
QY      846 -NSRGLSNARKFLVYSATNGDVPSCSTTIGQDADRRYSQNSLGEMTKLGSALTNRDPT 904
Db      973 SHGRG---DRELVVLSKENGSV-----WKHSRSGEYSTLDQILN-GMDEEGSLSE 1020
QY      905 YRRROMQRTITROMSFDTLKLVT---EDWFSDISPQMRRLNLTVSYGRRLRANQISF 961
Db      1021 LKKRRCRLIT--TDEPLFYVMSKLCQD-YDIIPEG-----GSLKSLVPLVQATF 1070
QY      962 MNDRLASWINTLEQMPYRTSWLILYEETEGIPDQMTLK-----TIYERISK- 1008
Db      1071 PENNAVTKRYKALQ-----AQPYDELVTKLGNQATFSPITYVEPRRRKE 1116
QY      1009 -----NIPTKDVEPLEIDGDIRNEVF-----LSRIPVLVARYK 1046
Db      1117 HRPGLRITLPSPSWDNPDSGEQDTSLRLCSYIGTDAQWEDITGTTKLVYANCA 1176
QY      1047 VELPCTVNLDPKRLREIADVRAARQISIGLAPPLLHSGPPRAPSQSPSCST 1106
Db      1177 NF---TTNV--SAREWLSQCPRTAIVANPATILYEL-----TAV 1211
QY      1107 SENGFPAGGVSPQHSSTYSQMGTPQHPFTNRGSGPAPGV--VLNSLVDAVCE-- 1161
Db      1212 PYMAFV-----IFAKMNDPREGRLRCYCMTDKDYDKTLQHE 1249
QY      1162 -----KLQIEGLDQSMUPCYCTTIKANINGRIACINDELKEMNMFGMHLPRS 1215
Db      1250 NFEVAVARSDILEVLSGL-----FAELSG-----NLVPYKAAQORSFHHQSPRE 1295
QY      1216 TVLENRMAESHVPPDPFLESSESSGPAPHEPARASHNE--LPHTELS----- 1263
Db      1236 NPLAM-----PAKVRDSREPGSLSFLKKAKMYEDTOHILHNTWPCAKG 1344
QY      1264 -----SQPYTINSEFELNTLGLDEGAPR-----HSNLSWQOSTRTPT-SL 1304
Db      1345 SGAEDRRRTPPLAARYSLSESTPGSLSGTEQAKMNAVISEHLGLSMARELQFVS 1404
QY      1305 SLSNDSQSSIEISKLTDVQAEYRDAYREYIAQMSQUL-----EGGPGSTTIGRSSPSTY 1360
Db      1405 EDIN-----RIRVENPRSLLEQSVALLNLTWIREG-----QANMENLY 1443
QY      1361 YMGSSSGGSIHSNLEQEKGDSEPKPDGRKSLMRKGVIDYSSSGVSTNDASPLDPT 1420
Db      1444 TALSDIRGEIVNMLESGRQSRNLKPD-----RRHTDRDYSLSPSQNGYSSL-- 1492
QY      1421 TEEDE-----KSDQSGSKL-----LPGKSSERSLSLTOTDLKLGSGLR 1460
Db      1493 ---QDELLSPALSGCALSPLRADQWNEVAVLDAIPLAATHDMLKSDMQVVASGLTP 1550
QY      1461 QKLPSDEDE---SGTESDNTPLLKDKRKAREKVERVPSPEHSAEPIRTFKAREYL 1517
Db      1551 SLVTAEDSLSCSKAEDSDAT-----GHEWKLLEGALSEPRGRP-----LGSLELV 1596
QY      1518 SDALLDKKDS-----DSGVASSSESPNHSIHLNEYADDSQLEKANLLELEDSSHGK 1569
Db      1597 EDDTVDSDATNGLIDLLQEEGGRSEKRLPGSKRODQATGAGQ--DSENEVSLVSGHORQ 1695
QY      1570 RGIFHSLSGLQDPIIARMSICEPK-----KSPSECSLIASSPENN 1611
Db      1656 ARITHS-----PIVSQVTERSQRLQDWDADGSIYSILODAAGSW 1696

RESULT 11
ID      08YTG9 PRELIMINARY: PRT: 426 AA.
AC      08YTG9:
DT      01-MAR-2002 (Tremblrel. 20, Created)
DT      01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT      01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE      Hypothetical protein A112748.
GN      A112748.

```

```

OS      Anabaena sp. (strain PCC 7120).
OC      Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX      NCBI_TaxID=103690;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=21595285; PubMed=11759840;
RA      Kaneo T., Nakamura Y., Molk C.P., Kuritz T., Sasamoto S.,
RA      Watanabe A., Iriuguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA      Kishida Y., Kohara M., Matsuno M., Matsuno A., Muraki A.,
RA      Nakazaki N., Shimpō S., Sugimoto M., Takazawa M., Yamada M.,
RA      Yasuda M., Tabata S.;
RT      *Complete genomic sequence of the filamentous nitrogen-fixing
RL      cyanobacterium Anabaena sp. strain PCC 7120.
RL      DNA Res. 8:205-213(2001).
DR      EMBL, AP003590; BAB74447.1; -
DR      InterPro: IPR002110; ANK.
DR      Pfam: PF00023; ank.13.
DR      PRINTS: PR01415; ANKYRIN.
DR      SMART: SM00248; ANK.13.
DR      PROSITE: PS50088; ANK_REPEAT.11.
DR      PROSITE: PS50297; ANK_REPEAT.1.
KM      Hypothetical protein; Complete proteome.
SQ      SEQUENCE 426 AA; 43800 MW; CF0B9E5D43ACAA03 CRC64;

Query Match          6.1%; Score 544; DB 16; Length 426;
Best Local Similarity 37.2%; Pred. No. 3.3e-26;
Matches 149; Conservative 64; Mismatches 180; Indels 8; Gaps 6;

QY      9 VINVEENIPALKALLECKDYDERNEGCGPTMTAABOGNLEIYKELIKNGANCNT-E 67
Db      8 LKVAKSGDIKGLGALLAGVGVDICDRGDTALFMANIGYEIVRSILDDGAVNYLAR 67
QY      68 DLONMTALISAKREGHVHYEELKGVNLNHHDMGCMTALMACYKGTDTVEYLLSNG 127
Db      68 KRYGLTALMAASANDVDIVQLTISRGAAVNATNEDSGTALAALAKGVNVAARYVLLAG 127
QY      128 ANPSVTGLQYVPIIWAAGRGHADIYHLLONGCAVNCSDKYGTPTVYMAARKHLECY 187
Db      128 ADVNITDKDDDT-ALKLAVKRGQAAVYQILLPSGADANEDDEGETLLMLADSGHDVY 186
QY      188 KILLMAGADVDGCAANSMTALYAVAGYGTQSYKEILKNPNVNTLTKDGNTALMASKE 247
Db      187 QVLLAAGVVDNEGNOGDTALLAAVAAGCAIAKILLDAGVADNHODQGESALHATVE 246
QY      248 GHTEIYODLDAGTYVINIDRSQDYLIGAVNGHVEIYRALLOKADIDINGQD-NKTA 306
Db      247 GYVDVYQVLLNGCANQIKNKLGDIPLVLAALQGHQIVETIL-KTG-ANVAGDNDGETP 304
QY      307 LYWAVERGNATVYRDILOCPNPTETCTKDETPLIKATKRNIEVELLDGKARVSAVD 366
Db      305 LTLAASQGHVATVRIILDDYGANNANIPASDGKTALIKATERNPVYQILLAGAVVNYQD 364
QY      367 KKGDTPLHIAIRSRKLAELLRNKDKDRLLYRPKAGET 407
Db      365 SVGATALLMAASGYNKVYQILLLEGADTNTL--KRRGGYT 402

RESULT 12
ID      061307 PRELIMINARY: PRT: 1943 AA.
AC      061307; 061305; 061306; 061308; 061309; 061310; 008866; 008867;
DT      01-JUN-1998 (Tremblrel. 06, Created)
DT      01-JUN-1998 (Tremblrel. 06, Last sequence update)
DT      01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE      Ankyrin 3 (Ankyrin G) (Epithelial ankyrin) (Ankyrin-3).
GN      ANK3.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A. (ISOFORMS 1-6).
SQ      STRAIN=C57BL/6J; TISSUE=KIDNEY;

```



```

Db 666 LLSRANANVL.SNKSGLTPLHLAEDRVNAEVLVNOGAHYDAQTCKMGTPLHNGCH--Y 723
QY 470 GSGK--SFLTK--KLEDEMT--FAGQO--IEPLFQPSMLIVLTLGGL 513
Db 724 GKIKIVNTHLOHSAVNAKNGYALHQAQOQTHIINVLONN----- 769
QY 514 GLLFAFVHPNLGIVSLFALLIFYFIVYFGGRGESMNMAWVLTSLRHIGYLE 573
Db 770 ASPNELTYNGNLTALM-----ARLGTIS 793
QY 574 LLLKMFV-----NPEPLPQTTKALPVRLEFDYNLSSVGGETSIAEMIA 620
Db 794 VYDTLKYVTEELMTTITTEKHKMNPETMNEVDL--MSDEYKAKASAPKLSOGEYI- 849
QY 621 TLDACERE-----EGF-LATR--LEFVFTED--T 646
Db 850 --SDGEEDATGTGTDKYLGPQDLKELGDSLPAGEYVGFSLARASALSRSFSDRSYT 907
QY 647 QGKKKKWTKCCLPSFVIFLFIIGCIISGILLALFRVDP-KHLT-----VAVLISI 697
QY 908 LNRSSYAR--DSMMIEELLVPSKEQHLTFTRFSDSLRHYSMAADTLDNVNLVSPV 963
QY 698 ASVGLAFVLCRTWQVYDLSLNSQKRLNMA-----KLKL- 737
Db 964 HSGFLVSPWDARG-----GSMGRSHHGMRIIPPRKCTAPRTICRLYKRRKLA 1014
QY 738 --KSEG-----FMVLKCEVELAMAKATIDSPFN 766
Db 1015 NPPVVEGEGLASRLVEMKPGAGQFLGPIVIEIPHGSMRKKEELLIVSENETWKEH 1074
QY 767 Q-----TLVYITIGLACEDQKVLQMDIVRV--LESGGFIAFASDPIIITKAINON 819
Db 1075 QFDSKNEDLABLNGMD--EELDSPEELGTKRICRIITK-DEPQYFA-----VYSRIKOE 1126
QY 820 LNSVLRDSNINHDMYRNIIVLPLVLSRGLSNARKFLV-TSAINGDVPCSDTIGQ--- 875
Db 1127 SNOIGPEGI-----LSSTTVPLVOASPEGALTKRIRVGLAQP 1166
QY 876 --EDARRVSONSLGEMTKIGSKTALNRDYY--RRROMOTITROMSFDLTJLKL---VT 928
Db 1167 VPETVKKI-----LGNKATFSPIVTVEPRRRKFKHPLT--MTIPVPPSGGVS 1214
QY 929 EDWSDSIPQMRRLNIVSYTGILLANQISFWMDLASVINLTEMPP-----YR 979
Db 1215 NGYGDAMP-NLRLLCSITGTSP-----AQMEDITGTTPTLTKDCVSFT 1259
QY 980 TS-----WL-----ILYLETEGIPDQMTLKI-YERISKNIPPTKVELELDGIR- 1027
Db 1260 TNVSARFMADCHQVL--ETVGLASQLYRELLCVPMAFVYFAKTNDP--VSSLRG 1313
QY 1028 -----NFEVFLSRTPVLAARDYKFLPCTVNLDPKLR--ELIADVRA 1068
Db 1314 FCMETDHDVDTLEQOENFEVYARSK-DIEVLEGGPIYVDCYGNLAPLTKGGQLVNFYS 1372
QY 1069 AREQISIGLAPPLPLHEGPAPPSGY-----SQPSVSSSIFSNPFGAGVVSQPHS 1123
Db 1373 FKEN-----RLPESIKIRDTSOEP--CGRLSF----- 1397
QY 1124 SYSGMTGPHPFYNRSGPAPGVVLLNSLNDVACEKLQJEGDOSMLPQYCTTKK 1183
Db 1398 -----LKEPKTKGFLPQIV----- 1412
QY 1184 ANINGRYLAOCNIDELKEMNNMFGMHLFRSVLEKRNASHVVEDPREFLSESSGPA 1243
Db 1413 CNINITLPAHKAKEKADRRROS-----FASIALKRRYS-----YLPEPMSPO 1454
QY 1244 PHGEPAR-----ASHNELPHTELSQTPYTLNPSPELNTLGLDGAAPHMSUS-- 1293
Db 1455 --SPCRTRIRAIYADHLGLSWTELARE--LNFVDELINQIVEN--PNSLISOS 1504
QY 1294 -----W--QSOTRTPSLNSQDSIEISKLTDKVOAE-----RDAYREYIAQM 1338
Db 1505 FMLLKWVTRBDKNATDALTSLTKINRDIYTLLEGPFLFDGNGISGRSFADENNVYH 1564

```

```

QY 1339 SQLEGGFG-----STTISGRSSPHSTYVMGSSGGSINS--NL 1375
Db 1565 DPVDGHPSPFOVELETMPGLWTYPPNPFOQDHFSDISIESPRTPSRLSDGLVPSQNI 1624
QY 1376 EQEKG-----DSS--EPKPDGKRSFLMKRGVYIDSSGSVSTNDASPLDPIIEDE 1425
Db 1625 EHPGPGPVVTAADTSLSDSKMD-----SVTVDPDAPLDL--VDESQ 1665
QY 1426 KSDQSGSKL-----LPQKSSERS--LFQTDLKIKSGSLRYOKLPDEDESGTE---- 1473
Db 1666 LKDLQCEQAQCAVAPGIPNDGRQAEPRLPQRYKMGSEQOEKKSQSPDEYVEDKAK 1725
QY 1474 -----ESDNPPLKDKDKRAEKVERVP-----KSPHSAB-PITFETKA 1513
Db 1726 SLFEDIQLEVEEAEMTEIDQGAMLRVORAEIAMSLSLAWONETPSSGLESPPAQ- 1781
QY 1514 KEYISDALDKKSDSGVRS-----ESSPHSLHNEVADDSQLEKANLIEL 1561
Db 1782 -RLTGLGLDRDDSSDQARDSTSYLTGEPKIEANGHHTA--EYIPRA--KAKPYFP 1835
QY 1562 EDDSHSGKRGIT-----PHSLGLDPIIARMSICSEDKKSPSECS--LIASSPEMWP 1612
Db 1836 ESQNDIGKOSIKENLKPXTHGCGRTPEPV-----SPLTAYOKSLETSKLVIEDAPKCPV 1891
QY 1613 ACQKAVYNRPSTVTLNNSNAPANANONPDMEGIRTSQVILRPSSPMTIIONEN 1672
Db 1892 VGKK--KMTRTYD-----GKARLNLOEBS-----STRSEPK--QGGG 1926
QY 1673 LKSMTRK 1679
Db 1927 YKVKTKK 1933

```

RESULT 13

```

ID 08VC68 PRELIMINARY; PRT; 1726 AA.
AC 08VC68;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Hypothetical 188.2 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL: BC021657; AAH21657.1; -
DR InterPro: IPR002110; ANK.
DR InterPro: IPR000488; Death.
DR InterPro: IPR000906; Z05.
DR Pfam: PF00023; ank. 24.
DR Pfam: PF00531; death. 1.
DR Pfam: PF00791; Z05. 1.
DR PRINTS: PR01415; ANKRYIN.
DR SMART: SM00248; ANK. 22.
DR SMART: SM00005; DEATH. 1.
DR SMART: SM00218; Z05. 1.
DR PROSITE: PS50088; ANK_REPEAT. 21.
DR PROSITE: PS50297; ANK_REPEAT_REGION. 1.
DR PROSITE: PS50017; DEATH_DOMAIN. 1.
KW Hypothetical protein.
SQ SEQUENCE 1726 AA; 188239 MW; 430148667E0B1CA CRC64;

```

Query Match

Best Local Similarity 5.9%; Score 522.5; DB 11; Length 1726;
Matches 406; Conservative 246; Mismatches 612; Indels 789; Gaps 82;

14 EENIPALKLLECKDQVDERNECGQTPMLTAEQNLLEYKLLIKNGANCNELDLDMWT 73


```

Db      1476 RGEIYNMLEGSGRGSRNLKPD-----RRHTDRDYSLSPSONMGRQGRQARITHSPTV 1527
QY      1418 DPTDEDEKS-----DQSGS-----KLBGKSSRSSTLFOYDLKLSGSL 1458
Db      1528 SOVTRSDRDLQDMADDSIYSYLQDAAGSGWQEVYGGPPHFGQSTMTREGLR-PGSGQ 1586
QY      1459 RYOKLPSEDESGTEESDNTPLLDKDKRAEGKVERYPKSPHSAEPITFTIKAEYLS 1518
Db      1587 EYEKLVSVSEHTMTQGEAPASSQADRROQGOGEVY-----QKAKNTFTQVYOGNEFON 1642
QY      1519 ---DALLDKRDSSSG-----VRSESSPPNHSL--HNEVADDSOLEKANLIELE 1562
Db      1643 IPGEQVYEEQFTDEQGNIVTKIIRKVVROIDLSSADAAQHEHEVTEGPLEDPSLEVD 1702
QY      1563 DD---SHS 1567
Db      1703 IDYFMKHS 1710

```

RESULT 15

Q12955

AC 012955: PRELIMINARY: PRT: 4377 AA.

```

DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Ankyrin G.
GN ANK-3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN 11
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN STEM;
RA MEDLINE=95138209; PubMed=7836469;
RT Kordell E., Lambert S., Bennett V.;
RT "Ankyrin, A new ankyrin gene with neural-specific isoforms localized
RT at the axonal initial segment and node of Ranvier.";
RL J. Biol. Chem. 270:2352-2359(1995).
RN 12
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN STEM;
RA Carpenter S.S.;
RL Submitted (Aug-1994) to the EMBL/GenBank/DBJ databases.
DR HSSP; P55273; 1818.
DR InterPro: IPR002110: ANK.
DR InterPro: IPR000488: Death.
DR InterPro: IPR000906; 205.
DR Pfam: PF00023; ank; 24.
DR Pfam: PF00531; death; 1.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 21.
DR SMART; SM00218; ZUS; 1.
DR PROSITE; PS50088; ANK_REPEAT; 21.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
KW ANK repeat; Repeat.
SQ SEQUENCE 4377 AA; 480395 MW; F42379E55768B684 CRC64;

```

Query Match 5.9%; Score 521; DB 4; Length 4377;
 Best Local Similarity 20.6%; Pred. No. 5, 1e-23;

Matches 427; Conservative 284; Mismatches 695; Indels 668; Gaps 87;

```

QY      14 EENIIPALAKLCKCDVDERNECGOTPLMIAEGNLEIVKELINGANGCNLELDLDMNT 73
Db      83 KEHVEVSELLQREANVDAAATKKGTALHIASLAGOAEVYKVLVTNGANVANVAQNGPFT 142
QY      74 ALISASKEGHVHIVEELKCGVNLHNRDMGWTALMAGCYKKGITVVELL----- 124

```

```

Db      143 PLVMAQENHLEVYKFLDNGASQSLATEDGFTPLVVALQGGDQVSLLENDTKGKVR 202
QY      125 -----SHGANSPTGL----- 135
Db      203 LPAHLIARKDPTKAALLLQNDNMADVESKSGFTPLHIAHGNIINVTLLINRAAVD 262
QY      136 ---QSYVPLTMAAGSHADIVALLQNGAKVCSKRYGTPPLVMAKRG----- 182
Db      263 FTARNDDTPLHAVASKRGNAMVRLDLDRGAKIDAKTRDGLTPLHCGARGSHQVYEMLD 322
QY      183 -----HLRCVKTLAMGADVOEGANSWTALIVAKGTYOS 219
Db      323 RAAPILSKTKNGSLPLMATQSGHLCQVOLLQHNVPVDVYNDYITLALHVAHCGHYAV 382
QY      220 VKELIKRNPVNLTDKQNTALMASKE----- 247
Db      383 AKVLLDKKAPNAKALNGFTPLHACKKRIKYMELLKHGASIQAVTESGLPIHVAAF 442
QY      248 -GHTEIVODLDAGTYVNIPDRSGDTVLGAVRGHVEIVRALLQKADIDRGDQDKTA 306
Db      443 MGHVNIYSQLMHHGASPNNTNWGETALHMAARSGOAEVRYLYVDQAEAKAKDQTP 502
QY      307 LYMAVEKGNATVTRDILQ--CNPD-----EICT 333
Db      503 LHISARLGRADIVQQLQOGASPNNAFTSGYTPPLHSAREGHEDVAFLIDHGASLSTT 562
QY      334 KGEFTPLKATKRNIEVELLDKGAQVSAVDKGDTPPLHAIIRGRSKLAELLRNPK 393
Db      563 KKGFTPLHVAAKGKLEAVANLLQKSASPDAGKSLGPLHVAAHADNGKVALLL-- 619
QY      394 DGLTLRPMKAGTEPNIDC--SHQSTLQI--FGAR-----HLSPTEDGM 438
Db      620 QGASPHAAKNGTTPPLHIAKKNMQDIATLLEYGADANAVTROGIAVTHLAQGHVDM 679
QY      439 LGDYLSALADILSEPTMQP-----PICYGLY 466
Db      680 VSLILGRNANVNLNKSGLTPLHAAQEDRVAVAEVLVQGHVDAQFMGTPLPHVGH 739
QY      467 AQMGSK--SELLK-KLEDEMTFPAQOIEPFPQSWLIVPLTLICGG---LGLPAF 519
Db      740 --YGNIKIVNFIQHSKAVNAKTNG--YTPHQAQ-----QGHTHINVLQON 785
QY      520 TVHPN-----LGIASISFLALTYFFIVIFGGRGEGSNMNAVISTRLARH-I 569
Db      786 NNSPNRLVYNGMTALGIARLGIYSVDLKIYV-----EETMTTITVEKHM 834
QY      570 GYIELLLKIMFVNPPL-----PEQTKALPVRFLETDYRLSVY-GETSIA----- 616
Db      835 NVPEYTNVEVILMSDDEVKRAKNAPE-----MLSDGEYISDVEGEDAMTGTDXYL 884
QY      617 --EMATLSDACEREGFLATRL-----FRYFKED--YQKKKKWKTCCLPSEVILF 666
Db      885 GPQDLKEIGDSDSLPAEGYMGFSIGARSASLSRFSRSSDRSYTLNRSSYAR---DSMMIHEL 940
QY      667 IIGCLISGITLLAIFRVPD-KHLT-----VNAVLSIASVGLAVLNCRTMWQYLD 717
Db      941 IYPSKQKHLLTFRREDSLSLRHYSWAADPLDNVNVYSSPIHSGFLVSTWVDKRG----- 994
QY      718 SLINSQKRPLHNAASKLRLKLGEGFMKYL---KCEVELMARMAKTIDFTQONQRLVVI 773
Db      995 ---GSMRGSRHHG-----MRIIIPRKCTAP-----TRICRLV-- 1025
QY      774 IGDLDACBDQKYLQMLDVTYRVLFSGKFPALFASPPIIILAIINONLSYLRDSINGHD 833
Db      1026 -----KHKHLAN-----PPH-----GERGJSSIRLVEMGPAPAQ 1055
QY      834 YMRN-IYHLPVPLNRSGLSNARKFLVTSATNGVCSPTTGIOEDADRVYQNSIGEXTK 892
Db      1056 FLGPVIVELPHRGSRG--KERELIVLASENCE-----YKHEQFD--SKNE--DLTE 1102
QY      893 L--GSKTALNRDQYRRQOMORTITRQMSFDLTLLVTEWFSDISIPQTMRLINIVSVT 950

```

```
DB 1103 LINGDELDSPDELGKRRICRITTKDP-----POYFAVVS--RIKQESNOIGPE 1150
QY 951 GRLLRAN-----QISFNMWDLASWINLEQMPYRTSWLILYEETEGIPDQMTLKIYER 1005
DB 1151 GGILSSTVPLVQASPEGALTRIRIVGLQ-----AQPVPEIYVKILGNK 1196
QY 1006 ISKNIPYTKDVEPLEIDGDIRNEFVLSKRTPLYVARDVKVFLPCTVNLDPKLEIAD 1065
DB 1197 AT-----PSPIVTEPRRRKFKPITMTIPV----- 1222
QY 1066 VRAAREQISIGLAVPLPLEHGPAPRAPSQCP--PSVCSSTFNGPAGVVSPOPHS 1123
DB 1223 -----PPSGEGVSNKGKDTTNNLLCSIT-----GGTSP----- 1254
QY 1124 SYISGTMGPQHPFYNRGSGPAPGVVLIN-----SLAVD-----AVCEKLOIEGLDOSM 1173
DB 1255 AQMEDITGTT-----PLTFIKDCVSFTTNVSARFWLADCHQVLETVGLATOL 1301
QY 1174 -----LPQYCTTIKKANINGRVLACNI-----DELKK--EMMNFQDMHLPFRSTVLEM 1220
DB 1302 YRELICVPMKFFVFAKMDPVESSLRCFCMTDDKVDKTLGOENFE-----VA 1352
QY 1221 RNAESHVVPEDPRFLESSESGPAPHGEPARA-----SHNELPHT-----ELSSQTPYT 1269
DB 1353 RSKDIEVLEKRPITY-DCYGNLAPLTKGQQLVFNFSYFKENRLPFSIKINDTSQEPGCR 1411
QY 1270 LNFSEELNTLGLDEGAPRHSNLSQOTRRTPLSLNSODSSIEISKULTDKVQAEYRD 1329
DB 1412 LSFLEKERTTGLPOTACNINITLPAKKETES-----DQDDIE--KTRDRQSFASL 1463
QY 1330 AYRE---YIAQMSOLEGGPGSSTTISGRSPHSITY-----YMGSSSGSGSIHSNLE 1376
DB 1464 ALRKRYSTLTERPGMIERTGAT-----RSLP--TYSYKFFSTRPYQSWTAPITV----- 1513
QY 1377 QEKGDSEPKPDDGKRSFLMRGDIVIDYSSSGVSTNDASPLDPI---TEDEKSDQSGS 1432
DB 1514 -----PGPAKSGF-----TSLSSSSSNTPSASPLKISWSVSTPSPIKSTLGAS 1556
QY 1433 KILPGKKSERSSLFQTLKLGSLR---YOKLPSDEDESGT-----EESDNPL----- 1480
DB 1557 TTSSVKSISDVASPIRS-LRTWSSPIKTVSOSQPYNIQVSSGTLARAPAVTEATPLKGLA 1615
QY 1481 -----LKDDKDKRAEGKVER-----VPKSP-----HSAEPIRPTI-KAKEYISDAL 1521
DB 1616 SMSTSSRTSPYTTAGLSLERSSITMTPPAPSKSNINMYSSSLPFXSITSAAPLISPL 1675
QY 1522 -----LDKDDSDSGVRSSESPNHSL--HNEVA---DDSOLEKAN--LIB-- 1560
DB 1676 KSVSPVKRSYDVVISAKITMASSLSLSPVKQMPGHAEVALYNGSISPLKYASSSTLNGC 1735
QY 1561 -----LEDDSHGKRGIPHSLGLODPIIARMSICSEDKSPSECSLIASSPEENWPACQ 1615
DB 1736 KATATLOEKISSATNSVSSVSAATDIYEKVFSTTAMPFSPLR-SIYSAAP----- 1786
QY 1616 KAYNLNRTPESTVTLNNSAPANRANQNFDEMEGIRETSOVILRPSSS-----PNPTIQN 1670
DB 1787 SAFQSLRTPPSASALYTSIGSSISATTS-----SVTSIITVPYVYVNVLPBPAALKL 1839
QY 1671 ENLAKMTKRGRSS--YTRLKSDPELHAAAS 1702
DB 1840 PDSNSFTKSAALLSPIKTLTETHPQPHFSRTS 1873
```

Search completed: July 1, 2003, 14:40:31
Job time : 112 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2003, 14:34:04 ; Search time 55.5 Seconds

(without alignments)
4117.564 Million cell updates/sec

Title: US-10-021-571-2

Perfect score: 8853
Sequence: 1 MSVLISQSVINVEENIPV.....ELHAASSTSGFGRRESTL 1715

Scoring table:
BLOSOM62
Gapop 10.0 , Gapext 0.5

Archived: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	8639.5	97.6	1762	23	AAU96841	Rat kidneys220 prot
2	8526	96.3	1763	23	AAU96824	Human polypeptide
3	8266	93.4	1715	22	AAU96893	Human polypeptide
4	8266	93.4	1715	22	AAU96840	Human polypeptide
5	8266	93.4	1715	22	AAU96840	Human polypeptide
6	2961.5	33.5	705	22	AAU96840	Human colon cancer
7	2782.5	31.4	551	22	AAU96840	Human death domain
8	2408	27.2	543	22	AAU96840	Human protein sequ
9	2376.5	26.8	1498	22	AAU96840	Drosophila melanog
10	2260	25.5	513	22	AAU96840	Human extracellular

11	2150	24.3	624	22	ABG08697	Novel human diago
12	614	6.9	120	22	AAU40811	Human polypeptide
13	599.5	6.8	129	22	AAU94785	Human protein sequ
14	549	6.2	109	21	AAU56277	Human secreted pro
15	538.5	6.0	1724	22	AAU20513	Protein of App rel
16	534.5	6.0	1872	22	AAU9160	Human protein Ssq
17	526	5.9	4274	22	ABG00972	Novel human diago
18	521	5.9	4386	22	ABG07375	Novel human diago
19	521	5.9	4397	22	ABG21944	Novel human diago
20	504	5.7	2443	22	ABG60521	Drosophila melanog
21	500.5	5.7	187	22	AAU40779	Human polypeptide
22	492.5	5.6	784	22	ABG62206	Drosophila melanog
23	471	5.3	2119	22	ABG64823	Drosophila melanog
24	467.5	5.3	1549	22	ABG58328	Drosophila melanog
25	467.5	5.3	1549	22	ABG67412	Drosophila melanog
26	466.5	5.3	1745	19	AAU76068	Full length ankyl
27	466.5	5.3	1745	19	AAU76068	D. immitis ankyl
28	466.5	5.3	1745	19	AAU76068	D. immitis ankyl
29	466.5	5.3	1745	21	AAU1589	Ankylin protein se
30	459	5.2	743	23	ABG97453	Novel human protei
31	450	5.1	1133	22	ABG08071	Novel human diago
32	449.5	5.1	1097	22	ABG21941	Novel human diago
33	449.5	5.1	1097	22	ABG21941	Novel human diago
34	448	5.1	378	22	AAU64963	Human secreted pro
35	441	5.0	435	22	AAU66710	Human cell growth
36	441	5.0	435	22	AAU93879	Human protein sequ
37	441	5.0	627	23	AAU17136	Human cancer cell
38	438	4.9	740	22	AAU01030	Human death domain
39	438	4.9	1569	22	AAU74360	Human cytoskeleton
40	434	4.9	1377	22	ABG08072	Novel human diago
41	432	4.9	1267	22	ABG22790	Novel human diago
42	431.5	4.9	1053	22	AAU28174	Novel human secret
43	429	4.8	705	22	AAU64962	Gene 24 human secr
44	423.5	4.8	705	22	AAU20496	Human secreted pro
45	421.5	4.8	2441	22	ABG62231	Drosophila melanog

ALIGNMENTS

RESULT 1
AAU96841
ID AAU96841 standard; Protein: 1762 AA.
XX
AC AAU96841;
XX
DT 30-JUL-2002 (first entry)
XX
DE Rat kidneys220 protein from Genbank AAG35185.
XX
DE
XX
XX Kidins220; Kinase D interacting substrate of 22kDa; cytosolic;
KW neuroprotective; gene therapy; protein kinase D; PDK; cancer; rat;
KW neurodegenerative disease; glioblastoma multiforme; prostate cancer.
XX
OS
XX Rattus sp.
XX
XX WO20020786-A2.
XX
XX 14-MAR-2002.
XX
XX 06-SEP-2001; 2001MO-GB03977.
XX
XX 06-SEP-2000; 2000US-230449P.
XX
XX (IMCR) IMPERIAL CANCER RES TECHNOLOGY LTD.
XX
XX Schiavo G, Iglesias T;
XX
XX WPI; 2002-371879/40.
XX
XX Polypeptide kinase D interacting substrate of 220 kDa and
XX polynucleotide; useful for identifying modulators useful in treating
XX cancer and neurodegenerative diseases
XX
XX PT

XX Disclosure: Fig 18; 228pp; English.

CC The invention relates to polypeptide comprising a kinase D interacting
CC substrate of 220 kDa (Kidins220) from rat and human or their
CC fragment, variant or fusion provided that the protein is not the
CC polypeptide encoded by Genbank Accession No. AB033076. Also included are
CC a polynucleotide encoding Kidins220 provided that polynucleotide is
CC not any one of the clones corresponding to the 61 Genbank Accession Nos.
CC given in the specification; a recombinant host cell containing the polynucleotide
CC the polynucleotide; an expression/replicable vector comprising
CC or vector; an anti-Kidins220 antibody (used in the preparation of
CC Kidins220); an agent (AI) which modulates activity of Kidins220;
CC protein kinase D (PKD) or interaction between PKD and Kidins220;
CC a mutant animal transgenic for Kidins220; the use of an agent capable of
CC detecting the expression of Kidins220 gene in manufacture of a diagnostic
CC reagent for diagnosing or prognosing cancer or for monitoring the
CC progression of cancer in a patient; a compound comprising a moiety that
CC selectively binds to Kidins220 or its variant and another moiety; and a
CC nucleic acid molecule encoding the compound. The antibody is useful for
CC modulating the activity of Kidins220, where the antibody prevents a
CC region of Kidins220 interacting from another protein, or reduces the
CC ability of Kidins220 to bind to adenosine triphosphate (ATP). The
CC Kidins220 polynucleotide is useful for identifying an agent which
CC modulates the promoter activity of the polynucleotide, and Kidins220 is
CC useful for identifying an agent which modulates activity of Kidins220,
CC protein kinase D (PKD) and the interaction between PKD with Kidins220.
CC Kidins220 or its polynucleotide is useful in medicine, e.g. using gene
CC therapy or for diagnosing or monitoring progression of cancer in a
CC patient. The agent is useful in manufacture of medicament, for use in
CC treatment of neurodegenerative disease, and the agent which inhibits
CC function of Kidins220 gene or its product is useful for treating cancer
CC which is glioblastoma multiforme or prostate cancer. The compound is
CC useful for imaging cancer in an individual, and for diagnosing or
CC prognosing, and also for treating an individual. The present
CC sequence represents Rat Kidins220 protein from Genbank accession number
CC AA035185.

XX Sequence 1762 AA;

Query Match 97.6%; Score 8639.5; DB 23; Length 1762;

Best Local Similarity 95.8%; Pred. No. 0;

Matches 1691; Conservative 6; Mismatches 15; Indels 53; Gaps 5;

QY 1 MSVLISOSVINYVEENIPALKALLERCKDVERNEGQPTMLAEOGVEIYKELLN 60
Db 1 MSVLISOSVINYVEENIPALKALLERCKDVERNEGQPTMLAEOGVEIYKELLN 60
QY 61 GANCNLEDDNMWTLAISKSGHITHYBELLSKASLEHRDMGWTALMAACYKGRDYY 120
Db 61 GANCNLEDDNMWTLAISKSGHITHYBELLSKASLEHRDMGWTALMAACYKGRDYY 120
QY 121 ELLLSHGANGSVTGLQTSYPTIWAAGRGHADIYHLLONGAKYNSDKGTTPYWAAR 180
Db 121 ELLLSHGANGSVTGLQTSYPTIWAAGRGHADIYHLLONGAKYNSDKGTTPYWAAR 180
QY 121 ELLLSHGANGSVTGLQTSYPTIWAAGRGHADIYHLLONGAKYNSDKGTTPYWAAR 179
Db 121 ELLLSHGANGSVTGLQTSYPTIWAAGRGHADIYHLLONGAKYNSDKGTTPYWAAR 179
QY 181 KGHLECYKHLIAGADVDEGANSMTALIAVAKGYQTSYKELLRPNVNLTDKOGNTA 240
Db 181 KGHLECYKHLIAGADVDEGANSMTALIAVAKGYQTSYKELLRPNVNLTDKOGNTA 240
QY 180 KGHLECYKHLIAGADVDEGANSMTALIAVAKGYQTSYKELLRPNVNLTDKOGNTA 239
Db 180 KGHLECYKHLIAGADVDEGANSMTALIAVAKGYQTSYKELLRPNVNLTDKOGNTA 239
QY 241 LMTASRGHTEIYODLLDAGTYVNIIPRSGDYVLIGAVRGHVEIYALLQKTAADIDIRG 300
Db 241 LMTASRGHTEIYODLLDAGTYVNIIPRSGDYVLIGAVRGHVEIYALLQKTAADIDIRG 300
QY 240 LMTASRGHTEIYODLLDAGTYVNIIPRSGDYVLIGAVRGHVEIYALLQKTAADIDIRG 299
Db 240 LMTASRGHTEIYODLLDAGTYVNIIPRSGDYVLIGAVRGHVEIYALLQKTAADIDIRG 299
QY 301 QDNKTALYMAVEKGNATMVDILOCNPDTEICTKDGSTPLIKATKMNIEVEVLLDKGA 360
Db 301 QDNKTALYMAVEKGNATMVDILOCNPDTEICTKDGSTPLIKATKMNIEVEVLLDKGA 360
QY 300 QDNKTALYMAVEKGNATMVDILOCNPDTEICTKDGSTPLIKATKMNIEVEVLLDKGA 359
Db 300 QDNKTALYMAVEKGNATMVDILOCNPDTEICTKDGSTPLIKATKMNIEVEVLLDKGA 359
QY 361 KVSADVKKGDTPLHVAIRGRSRRLAELLRNPKDGRLLYRPNKAGETPNYIDCSHOKSLI 420
Db 361 KVSADVKKGDTPLHVAIRGRSRRLAELLRNPKDGRLLYRPNKAGETPNYIDCSHOKSLI 420
QY 360 KVSADVKKGDTPLHVAIRGRSRRLAELLRNPKDGRLLYRPNKAGETPNYIDCSHOKSLI 419
Db 360 KVSADVKKGDTPLHVAIRGRSRRLAELLRNPKDGRLLYRPNKAGETPNYIDCSHOKSLI 419
QY 421 TQIFGARHLSPTIEDGMGLGYDLYSSALADIISEPTMOPPTICGLIYAQMGSGSFLTKLI 480
Db 421 TQIFGARHLSPTIEDGMGLGYDLYSSALADIISEPTMOPPTICGLIYAQMGSGSFLTKLI 480

Db 420 TQIFGARHLSPTIEDGMGLGYDLYSSALADIISEPTMOPPTICGLIYAQMGSGSFLTKLI 479
QY 481 EDEKKTFFAGQOTPELFQSWLIYFLLTLLCGGLGVFAFVPDNLAIASLFLAIYIF 540
Db 480 EDEKKTFFAGQOTPELFQSWLIYFLLTLLCGGLGVFAFVPDNLAIASLFLAIYIF 539
QY 541 FIYIYGGREGGSSWMMWMAALSTRLARHIGYBELFKLMFVNPPELPEOTYALPVRLF 600
Db 540 FIYIYGGREGGSSWMMWMAALSTRLARHIGYBELFKLMFVNPPELPEOTYALPVRLF 599
QY 601 TDNRRLSVGGESLAEIMATLSDACEREGFATLRFVFRRESGKKKKKTCCLPS 660
Db 600 TDNRRLSVGGESLAEIMATLSDACEREGFATLRFVFRRESGKKKKKTCCLPS 659
QY 661 FVIFLEFVIGCIAGITLLAIFRVPKHLTVNALISIASVGLAFVLCRTWMOYLDLSL 720
Db 660 FVIFLEFVIGCIAGITLLAIFRVPKHLTVNALISIASVGLAFVLCRTWMOYLDLSL 719
QY 721 NSQKRLHSAASKLHLKLSGFEKKVYKCEVELMARAKTIDSTQOTRLVYIDGLDAC 780
Db 720 NSQKRLHSAASKLHLKLSGFEKKVYKCEVELMARAKTIDSTQOTRLVYIDGLDAC 779
QY 781 EODKVLQMLDTYVLFSGKPELAFASDPHIIKAIKINOLNSYLKDSNINHDYRNIVH 840
Db 780 EODKVLQMLDTYVLFSGKPELAFASDPHIIKAIKINOLNSYLKDSNINHDYRNIVH 839
QY 841 LPTFLNSRGLSNARKLTVSATNGIDTSGDTGTGQEDTRVRSQNSLGEMTKLSTALN 900
Db 840 LPTFLNSRGLSNARKLTVSATNGIDTSGDTGTGQEDTRVRSQNSLGEMTKLSTALN 899
QY 901 RRTYRRRQORITRQMSFDLTKLTVTDWSDISPOMRRLNIVSTGLLAAQNT 960
Db 900 RRTYRRRQORITRQMSFDLTKLTVTDWSDISPOMRRLNIVSTGLLAAQNT 959
QY 961 FNMRLASWINLLEQMPRTSWLLIYLEETEGLPDQMTLKYERISKNIPTTKDVEPLL 1020
Db 960 FNMRLASWINLLEQMPRTSWLLIYLEETEGLPDQMTLKYERISKNIPTTKDVEPLL 1019
QY 1021 EIDGDIRNEFVFLSSTPVLVADYKTELPCTVNDLPKRELIADYRAAREINGGLAY 1080
Db 1020 EIDGDIRNEFVFLSSTPVLVADYKTELPCTVNDLPKRELIADYRAAREINGGLAY 1079
QY 1081 PPLPLEHGPPRPSPGSGQSPASVCSASFNGPPPGGVSPQPSYSGLSGQHPFYNR 1139
Db 1080 PPLPLEHGPPRPSPGSGQSPASVCSASFNGPPPGGVSPQPSYSGLSGQHPFYNR 1139
QY 1140 -----AAMPATGSSLL 1151
Db 1140 FFAPLYLTPRYYPGGSQHLISRSYKTSLPDQNNGLPCDSGFNKRQOAAVATGSSLL 1199
QY 1152 SSMTVVYVCEKRLQTEGLDQNNMPOCTTIKANKANGRLVSCNIDELKKEAMFGWMH 1211
Db 1200 SSMTVVYVCEKRLQTEGLDQNNMPOCTTIKANKANGRLVSCNIDELKKEAMFGWMH 1259
QY 1212 LFRSMVLEKRSVESQVYPPEDPRLNENSSAPVPHESARRSSHTLPTLSSQPTPLN 1271
Db 1260 LFRSMVLEKRSVESQVYPPEDPRLNENSSAPVPHESARRSSHTLPTLSSQPTPLN 1319
QY 1272 FSEFEELNTLGLDEGAPRHSNLSWSQOTRTPSLSSLSNODSSIEISKLTQVQAEYRDAY 1331
Db 1320 FSEFEELNTLGLDEGAPRHSNLSWSQOTRTPSLSSLSNODSSIEISKLTQVQAEYRDAY 1379
QY 1332 REYIAQMSQLEGGTSSITSGRSPHSHTYIYGQSSGGSHTSTLEOEGKGEKLOEGR 1391
Db 1380 REYIAQMSQLEGGTSSITSGRSPHSHTYIYGQSSGGSHTSTLEOEGKGEKLOEGR 1439
QY 1392 KSFLLKRGVDYIDSSSGVSTNEASPLDPTTEDEKSDSGSKLLPGKSSERPLFOQDL 1451
Db 1440 KSFLLKRGVDYIDSSSGVSTNEASPLDPTTEDEKSDSGSKLLPGKSSERPLFOQDL 1499
QY 1452 KIKGGGLYOKLPSDEDSGTRGVQITPHCSKMTKTKLAK -ORECASPOHSAEPIRT 1510
Db 1452 KIKGGGLYOKLPSDEDSGTRGVQITPHCSKMTKTKLAK -ORECASPOHSAEPIRT 1510

Db 1500 KIKGGLRYOKLPDEDESGTEESDNTF-LTKDKKAKGAERVKSP-EHSAEPIRT 1557
 QY 1511 FIKAKXYLSDALLDKKSSDGVSRNSENPNHSHNEADDSQLEKANIIELEDGHSK 1570
 Db 1558 FIKAKXYLSDALLDKKSSDGVSRNSENPNHSHNEADDSQLEKANIIELEDGHSK 1617
 QY 1571 RGMPHLSGLDPIIARMSICSEDKSPSECSLIASSPESEWPACOKANLNRPTSTYL 1630
 Db 1618 RGMPHLSGLDPIIARMSICSEDKSPSECSLIASSPESEWPACOKANLNRPTSTYL 1677
 QY 1631 NNNTAPTPNANONFDEIEGIRETSOYILRPGSPNPFAVONENLKSMAHRSQSSYTR 1690
 Db 1678 NNNTAPTPNANONFDEIEGIRETSOYILRPGSPNPFAVONENLKSMAHRSQSSYTR 1737
 QY 1691 SKDASELHAASSESTGFEGERESTL 1715
 Db 1738 SKDASELHAASSESTGFEGERESTL 1762
 RESULT 2
 ID AAU80244 standard; Protein: 1763 AA.
 AC AAU80244;
 DT 30-JUL-2002 (first entry)
 XX
 DE Rat Kidins220 protein.
 KW Kidins220; kinase D interacting substrate of 22kDa; cytosolic;
 KW neuroprotective; gene therapy; protein kinase D; PKD; cancer; rat;
 KW neurodegenerative disease; glioblastoma multiforme; prostate cancer.
 OS Rattus sp.
 XX
 PN WO200220786-A2.
 PD 14-MAR-2002.
 PF 06-SEP-2001: 2001MO-GB03977.
 PR 06-SEP-2000: 2000US-230449P.
 PA (IMCR) IMPERIAL CANCER RES TECHNOLOGY LTD.
 PI Schiavo G, Iglesias T;
 DR WPI: 2002-371879/40.
 XX N-PSDB: ABR51214.
 PT Polypeptide kinase D interacting substrate of 220 kDa and
 PT polynucleotide, useful for identifying modulators useful in treating
 cancer and neurodegenerative diseases
 PS Claim 33; Fig 1; 228pp; English.
 XX
 CC The invention relates to polypeptide comprising a kinase D interacting
 CC substrate of 220 kDa (Kidins220) from rat and human or the
 CC fragment, variant or fusion provided that the protein is not the
 CC polynucleotide encoded by Genbank Accession No. AB033076. Also included are
 CC a polynucleotide encoding Kidins220 provided that polynucleotide is
 CC not any one of the clones corresponding to the 61 Genbank Accession Nos.
 CC given in the specification; an expression/replacable vector comprising
 CC the polynucleotide; a recombinant host cell containing the polynucleotide
 CC or vector; an anti-Kidins220 antibody (used in the preparation of
 CC Kidins220); an agent (AI) which modulates activity of Kidins220;
 CC protein kinase D (PKD) or interaction between PKD and Kidins220;
 CC a mutant animal transgenic for Kidins220; the use of an agent capable of
 CC detecting the expression of Kidins220 gene in manufacture of a diagnostic
 CC reagent for diagnosing or prognosing cancer or for monitoring the
 CC progression of cancer in a patient; a compound comprising a moiety that
 CC selectively binds to Kidins220 or its variant and another moiety; and a
 CC nucleic acid molecule encoding the compound. The antibody is useful for

CC modulating the activity of Kidins220, where the antibody prevents a
 CC region of Kidins220 interacting from another protein, or reduces the
 CC ability of Kidins220 to bind to adenosine triphosphate (ATP). The
 CC Kidins220 polynucleotide is useful for identifying an agent which
 CC modulates the promoter activity of the polynucleotide, and Kidins220 is
 CC useful for identifying an agent which modulates activity of Kidins220,
 CC protein kinase D (PKD) and the interaction between PKD with Kidins220.
 CC Kidins220 or its polynucleotide is useful in medicine, e.g. using gene
 CC therapy or for diagnosing or monitoring progression of cancer in a
 CC patient. The agent is useful in manufacture of medicament, for use in
 CC treatment of neurodegenerative disease, and the agent which inhibits
 CC function of Kidins220 gene or its product is useful for treating cancer
 CC which is glioblastoma multiforme or prostate cancer. The compound is
 CC useful for imaging cancer in an individual, and for diagnosing or
 CC prognosing, and also for treating an individual. The present
 CC sequence represents Rat Kidins220.
 CC
 SQ Sequence 1763 AA;
 Query Match 96.3%; Score 8526; DB 23; Length 1763;
 Best Local Similarity 94.7%; Pred. No. 0;
 Matches 1671; Conservative 10; Mismatches 32; Indels 52; Gaps 6;
 QY 1 MSVLISQSVINVEEENIPALKALLEKCKDYDERNECGQPLMLAEOGVEIYKELLKN 60
 Db 1 MSVLISQSVINVEEENIPALKALLEKCKDYDERNECGQPLMLAEOGVEIYKELLKN 60
 QY 61 GANCLLELDLWMTALISAKSGHHIYBELLSKASLEHRMGWGTALMAACVGRDYY 120
 Db 61 GANCLLELDLWMTALISAKSGHHIYBELLSKASLEHRMGWGTALMAACVGRDYY 120
 QY 121 ELLISHGANPVTGLGYSVYPIIWAAGRGHADIYHLLONKACVNSDKYGTPLVVAAR 180
 Db 121 ELLISHGANPVTGLGYSVYPIIWAAGRGHADIYHLLONKACVNSDKYGTPLVVAAR 180
 QY 121 ELLISHGANPVTGLGYSVYPIIWAAGRGHADIYHLLONKACVNSDKYGTPLVVAAR 180
 Db 121 ELLISHGANPVTGLGYSVYPIIWAAGRGHADIYHLLONKACVNSDKYGTPLVVAAR 180
 QY 181 KGLECYVHLLAMGADVDOEGANSMTALIAVAKGYTQSYKEILRRPNVNLTDKDGNTA 240
 Db 181 KGLECYVHLLAMGADVDOEGANSMTALIAVAKGYTQSYKEILRRPNVNLTDKDGNTA 240
 QY 181 KGLECYVHLLAMGADVDOEGANSMTALIAVAKGYTQSYKEILRRPNVNLTDKDGNTA 240
 Db 181 KGLECYVHLLAMGADVDOEGANSMTALIAVAKGYTQSYKEILRRPNVNLTDKDGNTA 240
 QY 241 LMIASKEGHIETVDDLDAGTYVINIPDRSGTYVIGAVRGHVEIVALLQKADIDIRG 300
 Db 241 LMIASKEGHIETVDDLDAGTYVINIPDRSGTYVIGAVRGHVEIVALLQKADIDIRG 300
 QY 241 LMIASKEGHIETVDDLDAGTYVINIPDRSGTYVIGAVRGHVEIVALLQKADIDIRG 300
 Db 241 LMIASKEGHIETVDDLDAGTYVINIPDRSGTYVIGAVRGHVEIVALLQKADIDIRG 300
 QY 301 QDNKTALYWAEEKGATVWRDILQCNPTETCTKDGETPLIKATKMRNIEVEELLDKGA 360
 Db 301 QDNKTALYWAEEKGATVWRDILQCNPTETCTKDGETPLIKATKMRNIEVEELLDKGA 360
 QY 301 QDNKTALYWAEEKGATVWRDILQCNPTETCTKDGETPLIKATKMRNIEVEELLDKGA 360
 Db 301 QDNKTALYWAEEKGATVWRDILQCNPTETCTKDGETPLIKATKMRNIEVEELLDKGA 360
 QY 361 KVSADVKKGDTPLHVAIRGSRRLAELLRNPKDGRLLYRPNKAGETPYNIDCSHOKSIL 420
 Db 361 KVSADVKKGDTPLHVAIRGSRRLAELLRNPKDGRLLYRPNKAGETPYNIDCSHOKSIL 420
 QY 421 TQIFARHLSPTEIDGMDLYDLYSSALADILSEPTMQPCTCVGLYQWMSGKSFLLKTL 480
 Db 421 TQIFARHLSPTEIDGMDLYDLYSSALADILSEPTMQPCTCVGLYQWMSGKSFLLKTL 480
 QY 481 EDEMTFAGQOETPLFQFSWLVIFTLTLCGGLGVFAFVDVTLATAISFALYIF 540
 Db 481 EDEMTFAGQOETPLFQFSWLVIFTLTLCGGLGVFAFVDVTLATAISFALYIF 540
 QY 541 FIVIFEGRRGESESNMAMALSTRLARHIGYLELLFKIMFVNPELPBQTTKALPVRLF 600
 Db 541 FIVIFEGRRGESESNMAMALSTRLARHIGYLELLFKIMFVNPELPBQTTKALPVRLF 600
 QY 601 TDYNNLSVGGTSTLAEMIALTLSDACEREFPLATRLFRVTRTESOGKKKKKTCCLPS 660
 Db 601 TDYNNLSVGGTSTLAEMIALTLSDACEREFPLATRLFRVTRTESOGKKKKKTCCLPS 660
 QY 661 FVIFEFYIGCIAGITTLAIFRVDPKHLTVNAIILISAVGLAFVLCRTMVOVLDSL 720
 Db 661 FVIFEFYIGCIAGITTLAIFRVDPKHLTVNAIILISAVGLAFVLCRTMVOVLDSL 720
 QY 721 NSQRRRLSAASKLHLKSEGPMLVKCEVELMAMAKTIDSPQONQTRLVIIIDGLDAC 780
 Db 721 NSQRRRLSAASKLHLKSEGPMLVKCEVELMAMAKTIDSPQONQTRLVIIIDGLDAC 780

```

Db 721 NSORRKLHSAASKLHKLKSEGFMYKCEVELMAMAKTIDSFQNTQRLVYIIDGLDAC 780
QY 781 EDDKYLQMLDITVAVLEFSKPEFTAFASDPHIIIIKAIINQNSVLDSDNINGHDIYRNVH 840
Db 781 EDDKYLQMLDITVAVLEFSKPEFTAFASDPHIIIIKAIINQNSVLDSDNINGHDIYRNVH 840
QY 841 LVEFLNRSGLSNARKFLVTSATNGDITGDTGTGEDTRRVSQNSLSGEMFLGSKTALN 900
Db 841 LVEFLNRSGLSNARKFLVTSATNGDITGDTGTGEDTRRVSQNSLSGEMFLGSKTALN 900
QY 901 RRDYRRRQMQRTITRQMSFDLTKLLVTEWEDSDISPOTRMLNIVSYVGRLLRANOIT 960
Db 901 RRDYRRRQMQRTITRQMSFDLTKLLVTEWEDSDISPOTRMLNIVSYVGRLLRANOIT 960
QY 961 FMMDRLASWINLTEMQPYRTSWLIIYLETBGLPQOMLKTMYEISKNIPPTKVEPL 1020
Db 961 FMMDRLASWINLTEMQPYRTSWLIIYLETBGLPQOMLKTMYEISKNIPPTKVEPL 1020
QY 1021 EIDGDIRNEVEFLSSRTPLVAVADYKTFPLCTVNLDPKLTRETIADVRAAREQINIGLAY 1080
Db 1021 EIDGDIRNEVEFLSSRTPLVAVADYKTFPLCTVNLDPKLTRETIADVRAAREQINIGLAY 1080
QY 1081 PELPLHEGPPRPSPGSGYQPAVCSASFNGPFGVSPQPHSYSGLSGQPHFYNR- 1139
Db 1081 PELPLHEGPPRPSPGSGYQPAVCSASFNGPFGVSPQPHSYSGLSGQPHFYNR- 1139
QY 1140 -----AAVP-----ANGSSLLSSMTVD 1157
Db 1141 PPAPLYLPYPPYPGSQHLLISRSYKASLPDQNNGLPCDSGFKRQGSRCRHLQVLT 1200
QY 1158 VV-----CEKLRQIEGLDQNMPOYCTTIKKANINGHYLSOCNIDELKKEAMNFGDM 1211
Db 1201 EENDGRCREKLRQIEGLDQNMPOYCTTIKKANINGHYLSOCNIDELKKEAMNFGDM 1260
QY 1212 LERSVWLEBRVSQVPEDPFLFNENSGAPVPHGESARRSHHELPLTELSQPTPTLN 1271
Db 1261 LERSVWLEBRVSQVPEDPFLFNENSGAPVPHGESARRSHHELPLTELSQPTPTLN 1320
QY 1272 FSEFEELNTGLDEGAPRHSNLSMOSQTRTPSLSLNSODSSIEISKLTIDVOAEYRAY 1331
Db 1331 FSEFEELNTGLDEGAPRHSNLSMOSQTRTPSLSLNSODSSIEISKLTIDVOAEYRAY 1380
QY 1332 REYIAQMSQLEGSTISGRSPHSTYYIGQSSSGSIHSTLEQERKGELEKQEDGR 1391
Db 1381 REYIAQMSQLEGSTISGRSPHSTYYIGQSSSGSIHSTLEQERKGELEKQEDGR 1440
QY 1392 KSEFLMRGDIVIDSSGYSTNAPSPLDITEDEKSDSGSKLPGKSSSRPSPFQTDL 1451
Db 1441 KSEFLMRGDIVIDSSGYSTNAPSPLDITEDEKSDSGSKLPGKSSSRPSPFQTDL 1500
QY 1452 KLGKGLRYOKLPSDESGTGRVQITPHCSKMTIRTKRLKAK-OHECASPOHSAEPIRT 1510
Db 1501 KLGKGLRYOKLPSDESGTGRVQITPHCSKMTIRTKRLKAK-OHECASPOHSAEPIRT 1558
QY 1511 FIKAKYVLSDALDKKSDSGVRSNESPNNHSLHNEAADSOLEKANLLELEDEGHSGK 1570
Db 1559 FIKAKYVLSDALDKKSDSGVRSNESPNNHSLHNEAADSOLEKANLLELEDEGHSGK 1618
QY 1571 RGMPSLSGLDPIIARMSISEDKSPSECSLIASSPEEWPAQOKRYNNLRPSPVYTL 1630
Db 1619 RGMPSLSGLDPIIARMSISEDKSPSECSLIASSPEEWPAQOKRYNNLRPSPVYTL 1678
QY 1631 NNNTAPTRANQNFDEIGIRETSQVILRPSPNPTRAVQNNENKSMHAKRSQSSYTRL 1690
Db 1679 NNNTAPTRANQNFDEIGIRETSQVILRPSPNPTRAVQNNENKSMHAKRSQSSYTRL 1738
QY 1691 SKDASELHAASSESTGFEERESIL 1715
Db 1739 SKDASELHAASSESTGFEERESIL 1763

```

RESULT 3
AAM38993

```

ID AAM38993 standard; Protein: 1715 AA.
XX
AC AAM38993:
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 2138.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
OS Homo sapiens.
XX
PN MO20015312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000MO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI: 2001-442253/47.
XX
DR N-PSDB: AAI58149.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
XX such as central nervous system injuries -
XX
PS Example 4: SEQ ID NO 2138; 10078bp; English.
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AA42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localized neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemia and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 1715 AA.
XX
QY Query Match 93.4%; Score 8266; DB 22; Length 1715;
QY Best Local Similarity 92.7%; Pred. No. 0;
QY Matches 1591; Conservative 53; Mismatches 70; Indels 2; Gaps 2;
Db 1 MSVLISSQVINYVEENIPALAKLEKCDVDERNECGTPTMLAAEGQNVIVVELLN 60
1 MSVLISSQVINYVEENIPALAKLEKCDVDERNECGTPTMLAAEGQNVIVVELLN 60

```

```

QY 61 GANCNLEDLDMNTALISASKEGHIIVEELLKSGALSEHRDMGWTALMAGYKGTDDV 120
   |||
Db 61 GANCNLEDLDMNTALISASKEGHVHIVEELKCGVNLHRDMGWTALMAGYKGTDDV 120
QY 121 ELLSHGANSVTGLQYVYPIIWAAGRHADIVHLLONGAKVNCSDKYGTTPLVMAAR 180
   |||
Db 121 ELLSHGANSVTGLQYVYPIIWAAGRHADIVHLLONGAKVNCSDKYGTTPLVMAAR 180
QY 181 KGHLECYKHLLAGAVDQDGAANSKATLYAVKGGTQSYKELKRNPNVNLTDKQGNRA 240
   |||
Db 181 KGHLECYKHLLAGAVDQDGAANSKATLYAVKGGTQSYKELKRNPNVNLTDKQGNRA 240
QY 241 LMAASKEGHIEIYODLLDAGTYVNIPIRSGDVLIAVGRGHEIYRALLQKADIDING 300
   |||
Db 241 LMAASKEGHIEIYODLLDAGTYVNIPIRSGDVLIAVGRGHEIYRALLQKADIDING 300
QY 301 QDNKKTALYAVEKGNATWADILQCNPDIEICTKDETPLIKATKRNIEVEVLLDCKGA 360
   |||
Db 301 QDNKKTALYAVEKGNATWADILQCNPDIEICTKDETPLIKATKRNIEVEVLLDCKGA 360
QY 361 KVASAVDKKGDPTLHVAIRGRSRRLAELLRNPKDGLRLRNPKAGETPINIDCSHOKSIL 420
   |||
Db 361 KVASAVDKKGDPTLHVAIRGRSRRLAELLRNPKDGLRLRNPKAGETPINIDCSHOKSIL 420
QY 421 TQIFGARHLSPETDGLMGLYDLYSSALADILSEPTMOPICVGLYAQMGSGKSFLLKRL 480
   |||
Db 421 TQIFGARHLSPETDGLMGLYDLYSSALADILSEPTMOPICVGLYAQMGSGKSFLLKRL 480
QY 481 EDEMKTFAGQOTEPLOFQSWLIVFLLLCGGLVLPAPVDTNLAIAISLSFLALYIF 540
   |||
Db 481 EDEMKTFAGQOTEPLOFQSWLIVFLLLCGGLVLPAPVDTNLAIAISLSFLALYIF 540
QY 541 FIYIYGRGREGESMMNMAALSTRLARHIGYLELKLMPVNPPELOTKKLPVRFIF 600
   |||
Db 541 FIYIYGRGREGESMMNMAALSTRLARHIGYLELKLMPVNPPELOTKKLPVRFIF 600
QY 601 TDYNRLSSVGGETSLAEMIALTISDACEREFGLATRLFRVFRTEESOGKKMKKTCCLPS 660
   |||
Db 601 TDYNRLSSVGGETSLAEMIALTISDACEREFGLATRLFRVFRTEESOGKKMKKTCCLPS 660
QY 661 FVIFLFTVIGIINGITLALFRVDPKHLTVNATLITIASVYGLAFVINCNTMMQVLDLIL 720
   |||
Db 661 FVIFLFTVIGIINGITLALFRVDPKHLTVNATLITIASVYGLAFVINCNTMMQVLDLIL 720
QY 721 NSQKRLHSAASKLHLKLESGFMKVLKCEVELMARMAKTTDSTQNOTRLVLIIDGLDAC 780
   |||
Db 721 NSQKRLHSAASKLHLKLESGFMKVLKCEVELMARMAKTTDSTQNOTRLVLIIDGLDAC 780
QY 781 EODKVLQMDLTVRVLFSKGFPIAFASDPHIIKAINQNLNSVLRDSNINGHDMRYNIVH 840
   |||
Db 781 EODKVLQMDLTVRVLFSKGFPIAFASDPHIIKAINQNLNSVLRDSNINGHDMRYNIVH 840
QY 841 LPAVLNRSGLSNARKPLVTSATNGDITCSPTTQOETDPRVONSIGEMTKLGSSTALN 900
   |||
Db 841 LPAVLNRSGLSNARKPLVTSATNGDITCSPTTQOETDPRVONSIGEMTKLGSSTALN 900
QY 901 RRDYRRRQOMRTITROMSFDTLKLVTEDMFSDISPOTRRLNLTYSVGRLLRANQIT 960
   |||
Db 901 RRDYRRRQOMRTITROMSFDTLKLVTEDMFSDISPOTRRLNLTYSVGRLLRANQIT 960
QY 961 FNMORLASWINLTEOWPYRTSWILYLEETEGIPDQMTLKTIERISKINPTTKDVEPLL 1020
   |||
Db 961 FNMORLASWINLTEOWPYRTSWILYLEETEGIPDQMTLKTIERISKINPTTKDVEPLL 1020
QY 1021 EIIODIINFVFLSSRTPVLVARDVKTFLLPCTVNLDPKLEIILADVRAAREQISIGLAY 1080
   |||
Db 1021 EIIODIINFVFLSSRTPVLVARDVKTFLLPCTVNLDPKLEIILADVRAAREQISIGLAY 1080
QY 1081 PPLPLHAGPPRPSPGYSOPASVCSASFNGPPGCVVSPPHSSYSVGLSGPDPHPFNRA 1140
   |||
Db 1081 PPLPLHAGPPRPSPGYSOPASVCSASFNGPPGCVVSPPHSSYSVGLSGPDPHPFNRA 1140
QY 1141 AVPATGSSLLSSMTVDVCEKLRQIEGLDQNMMPQYCTTIKANINGRVLSCNIDEELK 1200

```

```

Db 1141 SGPAPGPVVLNSLVNDAVCEKLRQIEGLDQNMMPQYCTTIKANINGRVLSCNIDEELK 1200
QY 1201 KEMANFCDWHLFRSMVLKMSVSVQVVEPDPRLNENSAAPVPHGESARRSSTELPLT 1260
   |||
Db 1201 KEMANFCDWHLFRSTYLEMRNAESHVVEPDPREFLESSESSGPAPGPAPRASINELPHT 1260
QY 1261 ELSQTPYTLNFSPELNTLIGDEGAPRHSNLSMOSQTRRPTPSLSLKSQSSIEISLIT 1320
   |||
Db 1261 ELSQTPYTLNFSPELNTLIGDEGAPRHSNLSMOSQTRRPTPSLSLKSQSSIEISLIT 1320
QY 1321 DKVOAEYRDAYREYIAQMSOLEGGTSGSTISGRSSPHSTYVYIGOSSGSIHSTLEOERG 1380
   |||
Db 1321 DKVOAEYRDAYREYIAQMSOLEGGTSGSTISGRSSPHSTYVYIGOSSGSIHSTLEOERG 1380
QY 1381 KEGELKQEBGKRSFLMKRGDVIDYSSSGVTNESAFLDPIEEDPEKSPQSGSKILLPGKS 1440
   |||
Db 1381 KQSEKRPDDGRKSLMKRGDVIDYSSSGVTNESAFLDPIEEDPEKSPQSGSKILLPGKS 1440
QY 1441 SERPSLFQTDLKLKGGGLRYOKLPDEDESGTGRVQITPHCSKMITRRLKAKORECASP 1500
   |||
Db 1441 SERPSLFQTDLKLKGGGLRYOKLPDEDESGTGRVQITPHCSKMITRRLKAKORECASP 1500
QY 1501 QEHSAPITFTIKAREYLSDALLDKSSDSGVRNNESSPNHSLHNEAADDSOLEKANLI 1560
   |||
Db 1501 QEHSAPITFTIKAREYLSDALLDKSSDSGVRNNESSPNHSLHNEAADDSOLEKANLI 1560
QY 1561 ELEDGHSKGRMPHSLSGLDPIIARMSIGEDKKSSECSLASSPEESMPACOKAYN 1620
   |||
Db 1561 ELEDGHSKGRMPHSLSGLDPIIARMSIGEDKKSSECSLASSPEESMPACOKAYN 1620
QY 1621 LNRTPSTVTLNNTAPTRNANONFDEIGIRETSQVILRPGPSNPPTAVONENLKSMAHK 1680
   |||
Db 1621 LNRTPSTVTLNNTAPTRNANONFDEIGIRETSQVILRPGPSNPPTAVONENLKSMAHK 1680
QY 1681 RSORSSYTRLKSDASELH-AASSESTGGEEREESTL 1715
   |||
Db 1681 RSORSSYTRLKSDASELH-AASSESTGGEEREESTL 1715
QY 1680 RSORSSYTRLKSDPELHAAASSESTGGEEREESTL 1715
   |||
Db 1680 RSORSSYTRLKSDPELHAAASSESTGGEEREESTL 1715

RESULT 4
AAM39025
ID AAM39025 standard; Protein; 1715 AA.
XX
AC AAM39025;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 2170.
XX
KW Human; noctropic; immunosuppressant; cytosstatic; gene therapy; cancer;
peripheral nervous system; neuropathy; central nervous system; CNS;
Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
chemokine; thrombolytic; drug screening; arthritis; inflammation;
leukaemia.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000MO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.

```


QY 901 RDTYRRROMQRTITROMSFDLTLLVTEDEWSDISPOTRMLNIVSYGRLANQIT 960
 DB 901 RDTYRRROMQRTITROMSFDLTLLVTEDEWSDISPOTRMLNIVSYGRLANQIT 960
 QY 961 FMDRLASVINTEOMPYRTSWLLYLEETEGLPDMTLKTYERISKNIPTTKDVEPL 1020
 DB 961 FMDRLASVINTEOMPYRTSWLLYLEETEGLPDMTLKTYERISKNIPTTKDVEPL 1020
 QY 1021 EIDGDIRNEFVLSSTPVLVARDYKTLFCTVINDPKRELIARARARQINIGLAY 1080
 DB 1021 EIDGDIRNEFVLSSTPVLVARDYKTLFCTVINDPKRELIARARARQINIGLAY 1080
 QY 1081 PCLPLHEGPPRPSPQSPASVCSASFNGPPGVSPGVPSPHSPSYSGLSGPOHFYRRA 1140
 DB 1081 PCLPLHEGPPRPSPQSPASVCSASFNGPPGVSPGVPSPHSPSYSGLSGPOHFYRRA 1140
 QY 1141 AVPAAGSSLLSSMTVDVYCEKRLQIEGLDQNMFPQYCTTTKANINGRVLSQCNIDELK 1200
 DB 1141 AVPAAGSSLLSSMTVDVYCEKRLQIEGLDQNMFPQYCTTTKANINGRVLSQCNIDELK 1200
 QY 1201 KEMANFGDMHLFRSMVLEMRSEVQVYEDPRFENMSAPVPHGESARSSHTEPLT 1260
 DB 1201 KEMANFGDMHLFRSMVLEMRSEVQVYEDPRFENMSAPVPHGESARSSHTEPLT 1260
 QY 1261 ELSSOTPYTLNFSPEELNTLGLDEGAPRHSNLSMOSQTRRPPSLSSLSNODSIEISKLT 1320
 DB 1261 ELSSOTPYTLNFSPEELNTLGLDEGAPRHSNLSMOSQTRRPPSLSSLSNODSIEISKLT 1320
 QY 1321 DKVOAEYRDVAREYIAOMSOLEGGSSSTISGRSSPHSTYTYIGSSSGGSHSTLEDRG 1380
 DB 1321 DKVOAEYRDVAREYIAOMSOLEGGSSSTISGRSSPHSTYTYIGSSSGGSHSTLEDRG 1380
 QY 1381 KEGELKQEDGRKSLFKRGVDVITYSSGVSTNEASPLDPIEDDEKSPQSGSKLLPGKKS 1440
 DB 1381 KEGELKQEDGRKSLFKRGVDVITYSSGVSTNEASPLDPIEDDEKSPQSGSKLLPGKKS 1440
 QY 1441 SERPSEFOTDLKACGGLRYOKLPSEDEDESGTGRVOITPHCSKMITRKLAKORECASP 1500
 DB 1441 SERPSEFOTDLKACGGLRYOKLPSEDEDESGTGRVOITPHCSKMITRKLAKORECASP 1500
 QY 1501 GHSAPERTFTKAEVYSDALLDOKSDSGVRSNESPMSHSHNEAADSOLEKANLI 1560
 DB 1501 GHSAPERTFTKAEVYSDALLDOKSDSGVRSNESPMSHSHNEAADSOLEKANLI 1560
 QY 1561 ELEDGSHSGKRCMPHSLSGLODPIIARMSICSEDKSPSECSLIASSPEESMPACOKAYN 1620
 DB 1561 ELEDGSHSGKRCMPHSLSGLODPIIARMSICSEDKSPSECSLIASSPEESMPACOKAYN 1620
 QY 1621 LNRTSYTYLNNNTAFTIRANONDEIGIRETSQVILIRPGSPNPPIAVONENLKSMAHK 1680
 DB 1621 LNRTSYTYLNNNTAFTIRANONDEIGIRETSQVILIRPGSPNPPIAVONENLKSMAHK 1680
 QY 1680 RSORSSYTRLKSDASELH-AAASESTGGEERESTL 1715
 DB 1680 RSORSSYTRLKSDASELH-AAASESTGGEERESTL 1715
 RESULT 6
 AAG75604
 ID AAG75604 standard; Protein; 705 AA.
 AC AAG75604;
 DT 03-SEP-2001 (first entry)
 DE Human colon cancer antigen protein SEQ ID NO:6368.
 XX Human; colon cancer; colon cancer antigen; diagnosis; detection;
 KM colorectal carcinoma.
 OS Homo sapiens.
 XX

PN W0200122920-A2.
 PD 05-APR-2001.
 XX 28-SEP-2000; 2000WO-0526524.
 PF 29-SEP-1999; 99US-0157137.
 PR 03-NOV-1999; 99US-0163280.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA Ruben SM, Barash SC, Birse CE, Rosen CA;
 PI WPI: 2001-235357/24.
 DR N-PSDB: AAH35009.
 DR Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
 PT useful for preventing, diagnosing and/or treating colorectal cancers -
 PS Claim 11; Page 7830-7834; 9803pp; English.
 CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
 CC cancer-associated nucleic acid molecules (N) and proteins (P), where
 CC the proteins are collectively known as colon cancer antigens. The colon
 CC cancer antigens have cytostatic activity and can be used in gene
 CC therapy and vaccine production. N and P may be used in the prevention,
 CC diagnosis and treatment of diseases associated with inappropriate P
 CC expression. For example, N and P may be used to treat disorders
 CC associated with decreased expression by rectifying mutations or deletions
 CC in a patient's genome that affect the activity of P by expressing
 CC inactive proteins or to supplement the patient's own production of P.
 CC Additionally, N may be used to produce the colon cancer-associated ps,
 CC by inserting the nucleic acids into a host cell and culturing the cell
 CC to express the proteins. N and P can be used in the prevention, diagnosis
 CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
 CC and AAB77789 represent sequences used in the exemplification of the
 CC present invention.
 CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
 CC missing at time of publication, meaning no sequences are present for
 CC SEQ ID NO:1027 to 1052, 7921 and 7922.
 XX
 SO Sequence 705 AA:
 Query Match 33.5%; Score 2961.5; DB 22; Length 705;
 Best Local Similarity 95.5%; Pred. No. 3.5e-224;
 Matches 567; Conservative 12; Mismatches 14; Indels 1; Gaps 1;
 QY 62 ANCNEDLDNMTALLISAKREGHIIYELIKSGASLEHRDMGWTALMAYCKGRTDVE 121
 DB 28 SNCNEDLDNMTALLISAKREGHIIYELIKSGVNLHRDMGWTALMAYCKGRTDVE 87
 QY 122 LLLSHGANSPTGLQSYPIIWAAGRHADYHLLONGAVYNSDKGTPIVMAARK 181
 DB 88 LLLSHGANSPTGLQSYPIIWAAGRHADYHLLONGAVYNSDKGTPIVMAARK 146
 QY 182 GHLECYKHLIAMDADVDGANSMTALIVAVKGGTYQSVKEILKRNPNVLTDKDGNAL 241
 DB 147 GHLECYKHLIAMDADVDGANSMTALIVAVKGGTYQSVKEILKRNPNVLTDKDGNAL 206
 QY 242 MIAKSGHEIYODLDAGTYVNIIPRSGDYVLIGAVRGHVEIYRALLQKADIDIRQ 301
 DB 207 MIAKSGHEIYODLDAGTYVNIIPRSGDYVLIGAVRGHVEIYRALLQKADIDIRQ 266
 QY 302 DNKTALYMAVEKGNATMVDILOCNDPTEICTKDGTEPLIKATKKRNIEVVELLDKGA 361
 DB 267 DNKTALYMAVEKGNATMVDILOCNDPTEICTKDGTEPLIKATKKRNIEVVELLDKGA 326
 QY 362 VSAVDKKGDTPLHVALIGRSRRLABELLNNPRDGRLLYRPNKAGETPYNIIDCSHOKSILT 421
 DB 327 VSAVDKKGDTPLHVALIGRSRRLABELLNNPRDGRLLYRPNKAGETPYNIIDCSHOKSILT 386
 QY 422 QIFGARHLSPTETDGMGLGYLYSSALADILSEPTMOPPICVGLYLAOMSGSKFLKLLE 481

Db 387 QIFGARHLSPETEDGMDLYSSALADILSEPTMOPICVLGYAOWMGSGSKFLKLE 446
QY 482 DEKTRAGQOTEPFLQPSWLVFLTLGGLGLVAFVPTDNLALAIISFLALYIF 541
Db 447 DEKTRAGQOIEPLFOFSWLVFLTLGGLGLVAFVPTDNLALAIISFLALYIF 506
QY 542 IVIFGGRGSGESMNNAMALSTRARHIGYLELFLIMVNPPELEOFTKALPVFLPT 601
Db 507 IVIFGGRGSGESMNNAMALSTRARHIGYLELFLIMVNPPELEOFTKALPVFLPT 566
QY 602 DYNRLSSVGETSLAEMIALSDACEREGFLATRLFRVFRTEESOGKKRKKKT 655
Db 567 DYNRLSSVGETSLAEMIALSDACEREGFLATRLFRVFRTEESOGKKRKKKT 620
RESULT 7
ID AAE01035 standard; Protein: 551 AA.
AAE01035;
04-JUL-2001 (first entry)
Human death domain-containing receptor (DDCR) protein from HODDX59 clone.
Human: death domain-containing receptor; DDCR; immunosuppressive;
antiarthritic; antirheumatic; antiproliferative; cytostatic;
cardiac; vasotropic; cerebroprotective; neurotropic; neuroprotective;
antibacterial; virucide; fungicide; ophthalmological; gene therapy;
immunodeficiency disease; Acquired immune deficiency syndrome;
AIDS; leukaemia; autoimmune disease; systemic lupus erythematosus;
hyperproliferative disorder; neoplasm; cerebrovascular disorder;
cerebral ischaemia; angiogenesis; cardiovascular disorder;
neurodegenerative disease; Alzheimer's disease; Parkinson's disease;
ocular disorder; corneal infection; degenerative disease; SMA; apoptosis;
spinal muscular atrophy; epithelial cell proliferation; infection;
cancer; wound healing; skin aging; chemotaxis; HODDX59 clone.
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FT Domain 399..415 /label=Transmembrane_domain
FT Domain 426..442 /label=Transmembrane_domain
FT Region 64..69 /label=Immunogenic_epitope
FT Region 131..136 /label=Immunogenic_epitope
FT Region 164..169 /label=Immunogenic_epitope
FT Region 197..203 /label=Immunogenic_epitope
FT Region 229..235 /label=Immunogenic_epitope
FT Region 298..316 /label=Immunogenic_epitope
FT Region 328..335 /label=Immunogenic_epitope
FT Region 445..452 /label=Immunogenic_epitope
FT Region 482..490 /label=Immunogenic_epitope
FT Region 540..551 /label=Immunogenic_epitope
FT Region /label=Immunogenic_epitope
XX WO200129063-A2.
XX 26-APR-2001.
XX 17-OCT-2000; 2000WO-US28666.
XX 18-OCT-1999; 99US-0159585.
PR

PR 24-NOV-1999; 99US-0167246.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX NI J, Ruben SM;
XX MPI: 2001-300316/31.
XX N-PSDB: AAD04776.
XX
XX New death domain containing receptor polynucleotides and polypeptides,
XX useful for treating and diagnosing cancer
XX
XX Claim 1; Page 292-294; 298pp; English.
XX
XX The patent discloses novel death domain-containing receptor (DDCR)
XX cDNAs and their corresponding proteins. DDCR cDNA and protein are
XX used to prevent, treat or ameliorate a medical condition in mammals.
XX They are also used in diagnosing a pathological condition or
XX susceptibility to a pathological condition. The DDCR protein and
XX its antibodies are used in the diagnosis and treatment of disorders
XX such as immunodeficiency diseases (e.g. Acquired immune deficiency
XX syndrome (AIDS), leukaemia) autoimmune diseases (e.g. systemic lupus
XX erythematosus, rheumatoid arthritis), hyperproliferative disorders
XX (e.g. neoplasms of the breast or liver), cerebrovascular disorders
XX (e.g. cerebral ischaemia, angiogenesis), cardiovascular disorders
XX (e.g. cardiac arrest), neurodegenerative diseases (e.g. Alzheimer's
XX disease, Parkinson's disease), ocular disorders (e.g. corneal
XX infection), degenerative diseases (e.g. spinal muscular atrophy-SMA),
XX cancer, aberrant apoptosis, disorders of the placenta or uterus and
XX infections caused by bacteria, viruses and fungi. The DDCR proteins
XX are used to aid wound healing and epithelial cell proliferation,
XX to prevent skin aging due to sunburn, to maintain organs before
XX transplantation, to support cell culture of primary tissues, to
XX regenerate tissues and in chemotaxis. The DDCR sequences are used
XX in correcting aberrant cellular apoptosis by gene therapy.
XX The present sequence is human death domain-containing receptor
XX (DDCR) protein from clone HODDX59.
XX
XX Sequence 551 AA:
SQ
Query Match 31.4%; Score 2782.5; DB 22; Length 551;
Best Local Similarity 96.7%; Pred. No. 3e-210; 9; Indels 1; Gaps 1;
Matches 534; Conservative 8; Mismatches 9;
QY 102 MGSWTALMAMACYGRTDVVELLSHGANGPSYGLQSYVPIIIMAGRGHADVHLLONG 161
Db 1 MGSWTALMAMACYGRTDVVELLSHGANGPSYGLQSYVPIIIMAGRGHADVHLLONG 59
QY 162 AKVNSDKYGTPLVMAARKGHLECYKHLLAMGADVDGANSMTALIVAVGGYTOSYK 221
Db 60 AKVNSDKYGTPLVMAARKGHLECYKHLLAMGADVDGANSMTALIVAVGGYTOSYK 119
QY 222 EILKRPNNVLDKDSNTALMTASKRGHEIYODLLDATTYVNIIPRSGDYLICAVRG 281
Db 120 EILKRPNNVLDKDSNTALMTASKRGHEIYODLLDATTYVNIIPRSGDYLICAVRG 179
QY 282 HVEIVALLQKYADIDIRGODNKTALYMAVERGNATVBDILQCNPDTEICTKDETPLI 341
Db 180 HVEIVALLQKYADIDIRGODNKTALYMAVERGNATVBDILQCNPDTEICTKDETPLI 239
QY 342 KATKMRNIEVELLDKGAQVSAVDKGTPLHVALRGSRLAELLRNPDGRILYRP 401
Db 240 KATKMRNIEVELLDKGAQVSAVDKGTPLHVALRGSRLAELLRNPDGRILYRP 299
QY 402 NKAGETPVNIDCSHOKSILTOIFGARHLSPETEDGMDLYSSALADILSEPTMOPPI 461
Db 300 NKAGETPVNIDCSHOKSILTOIFGARHLSPETEDGMDLYSSALADILSEPTMOPPI 359
QY 462 CVGLGYAOWMGSGSKFLKLEDEKTRAGQOIEPLFOFSWLVFLTLGGLGLVAFV 521
Db 360 CVGLGYAOWMGSGSKFLKLEDEKTRAGQOIEPLFOFSWLVFLTLGGLGLVAFV 419
QY 522 DTNLALAIISFLALYIFFIYIYIFGGRGSGESMNNAMALSTRARHIGYLELFLKLMV 581
PR

DB 420 HNNLGIWVSLSEFLLIYIFIVYFGRREGESMMWAWLSPRLARHIGYLELLKLMFV 479
 OY 582 NPEPEDEOTKALPVAFLETDYNNRLSSVGGESTSLAEMTATLSDAECREGFATLRFV 641
 DB 480 NPEPEDEOTKALPVAFLETDYNNRLSSVGGESTSLAEMTATLSDAECREGFATLRFV 539
 OY 642 RTEESOGKKKK 653
 DB 540 KTEEDGKKKK 551

RESULT 8
 AAB95191
 ID AAB95191 standard; Protein; 543 AA.
 XX AAB95191;
 AC 26-JUN-2001 (first entry)
 XX
 DT Human protein sequence SEQ ID NO:17273.
 XX
 DE Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 XX Homo sapiens.
 OS
 XX
 PN EPI074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-0116126.
 XX
 PR 29-JUL-1999; 99JP-0248036.
 XX
 PR 27-AUG-1999; 99JP-0300253.
 XX
 PR 11-JAN-2000; 2000JP-0118776.
 XX
 PR 02-MAY-2000; 2000JP-0183767.
 XX
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 PA (HELI-) HELIX RES INST.
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 DR WPI; 2001-318749/34.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 XX
 Claim 8; SEQ ID 17273; 2537bp + CD ROM; English.

The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide which comprises a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any special methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification

CC of the present invention.
 XX
 SQ Sequence 543 AA;
 Query Match 27.2%; Score 2408; DB 22; Length 543;
 Best Local Similarity 86.8%; Pred. No. 1e-180;
 Matches 472; Conservative 24; Mismatches 46; Indels 2; Gaps 2;

OY 1173 MNPYCTTICKANINGVLSQCNIDELKEMANFGDMHFRSMVLEMSVEQVPEDP 1232
 DB 1 MNPYCTTICKANINGVLSQCNIDELKEMANFGDMHFRSMVLEMSVEQVPEDP 60
 OY 1233 RFLNENSSAVYPHGESARSSHTLPLTELSSQTPYTLNFSFELNTLGLDEAPRHSNL 1292
 DB 61 RFLNENSSAVYPHGESARSSHTLPLTELSSQTPYTLNFSFELNTLGLDEAPRHSNL 120
 OY 1293 SWSQTRRTPSLSSLNQDSSEISLTKTDVQVQAEYRDVAREYIAQMSQLEGTSSTISG 1352
 DB 121 SWSQTRRTPSLSSLNQDSSEISLTKTDVQVQAEYRDVAREYIAQMSQLEGTSSTISG 180
 OY 1353 RSPHSTYYIGOSSSGSISHTLEOERKGEELKOEGRKSPFLMKRGDVIDYSSGVSTN 1412
 DB 181 RSPHSTYYIGOSSSGSISHTLEOERKGEELKOEGRKSPFLMKRGDVIDYSSGVSTN 240
 OY 1413 EASPLDPTFEDEKSDQSGSKLLPGKKSSEPSLFTQDLKLGGLRYOKLPDEDESGT 1472
 DB 241 EASPLDPTFEDEKSDQSGSKLLPGKKSSEPSLFTQDLKLGGLRYOKLPDEDESGT 300
 OY 1473 GRVQITPHCSKMTKTKLAKKORCASPOHSAEPIRTFTKAEYISDALDKKSDSG 1532
 DB 301 EESDNTF-LTKDKRKAEKVERVPSHSAEPIRTFTKAEYISDALDKKSDSG 359
 OY 1533 VRSNESPNSHLNEADDSQLEKANLIELEDGSHGKGMPSLSGLDPTIARMSICS 1592
 DB 360 VRSNESPNSHLNEADDSQLEKANLIELEDGSHGKGMPSLSGLDPTIARMSICS 419
 OY 1593 EDKSPSECSLIASSPEESPACQKAYNLRRPSTYTLNNTAPTRANQFDELIGIRE 1652
 DB 420 EDKSPSECSLIASSPEESPACQKAYNLRRPSTYTLNNTAPTRANQFDELIGIRE 479
 OY 1653 TSOVYILRPGSPNPVAVONNLKSMARKRSQSSYTRLSKDSALH-AASSESTGGEER 1711
 DB 480 TSOVYILRPGSPNPVAVONNLKSMARKRSQSSYTRLSKDSALH-AASSESTGGEER 539
 OY 1712 ESTL 1715
 DB 540 ESTL 543

RESULT 9
 ABB64857
 ID ABB64857 standard; Protein; 1498 AA.
 XX
 AC ABB64857;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 21363.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 PN W0200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 XX
 PR 11-JUL-2000; 2000US-0614150.
 XX

KW inflammatory disorder; acquired immune deficiency syndrome; AIDS;
 KW Addison's disease; allergy; anaemia; asthma; atherosclerosis; melanoma;
 KW Crohn's disease; diabetes mellitus; atopic dermatitis; lymphoma; cancer;
 KW glomerulonephritis; multiple sclerosis; Grave's disease; osteoarthritis;
 KW osteoporosis; psoriasis; rheumatoid arthritis; ulcerative colitis;
 KW infection; cell proliferative disorder; actinic keratosis; myeloma;
 KW arteriosclerosis; neoplastic; anticonvulsant; antithyroid; nephrotropic;
 KW neuroprotective; dermatological.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Region 94..108 /note="Adrenocorticotrophin receptor"
 FT Domain 196..198 /note="Rgd cell interaction motif"
 FT Region 328..335 /note="Phage integrase family"
 FT
 PN WO200142285-A2.
 XX
 XX 14-JUN-2001.
 PF 05-DEC-2000; 2000WO-US32990.
 PR 10-DEC-1999; 99US-0172852.
 PR 16-DEC-1999; 99US-0172354.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 PI Yue H, Tang YT, Lal P, Burford N, Azimzal Y, Patterson C;
 PI Baughn MR, Lu DM, Shah P, Au-Young J;
 XX
 DR MPI: 2001-381632/40.
 DR N-PSDB: AAD08053.
 PT New human extracellular matrix and cell adhesion molecules and
 PT polynucleotide sequences encoding them, useful for diagnosis,
 PT prevention, treatment of genetic, autoimmune and cell proliferative
 PT disorders.
 XX
 PS Claim 1; Page 101-102; 135pp; English.
 XX
 CC The present sequence is a human extracellular matrix and cell
 CC adhesion molecule (XMD). The XMD is used for screening a compound for
 CC effectiveness as an agonist or antagonist of XMD. The identified agonist
 CC or antagonist are used for treating a disease or condition associated
 CC with decreased or increased expression of functional XMD. The
 CC polynucleotides encoding XMD are useful in somatic or germ-line gene
 CC therapy to correct a genetic deficiency, to express a conditionally
 CC lethal gene product and to express a protein which affords protection
 CC against intracellular parasites and also for diagnosis of disorders
 CC associated with expression of XMD. They are also used for generating
 CC hybridisation probes useful in mapping the naturally occurring genomic
 CC sequences and to create knock in humanised animals (pigs) or transgenic
 CC animals (mice or rats) to model human diseases. Oligonucleotide or longer
 CC fragments derived from the polynucleotide sequences may be used as
 CC elements on a microarray. Antibodies which specifically bind XMD may be
 CC used for the diagnosis of disorders associated with the expression of
 CC XMD, or in assays to monitor patients being treated with XMD. Diseases
 CC diagnosed, prevented or treated include genetic disorders such as
 CC adrenoleukodystrophy, Down's syndrome, cystic fibrosis, Gaucher's
 CC disease, myotonic dystrophy, sickle cell anaemia, thalassemia,
 CC autoimmune/inflammatory disorders such as acquired immune deficiency
 CC syndrome (AIDS), Addison's disease, allergies, anaemia, asthma,
 CC atherosclerosis, Crohn's disease, diabetes mellitus, atopic dermatitis,
 CC glomerulonephritis, multiple sclerosis, Grave's disease, osteoarthritis,
 CC osteoporosis, psoriasis, rheumatoid arthritis, ulcerative colitis,
 CC bacterial, fungal, parasitic, protozoal and helminthic infections and
 CC cell proliferative disorders such as actinic keratosis, arteriosclerosis
 CC and cancer including breast, bladder, bone marrow, brain and uterus
 CC cancer, leukaemia, adenocarcinoma, lymphoma, melanoma and myeloma.
 XX

SO Sequence 513 AA;
 Query Match 25.5%; Score 2260; DB 22; Length 513;
 Best Local Similarity 86.6%; Pred. No. 4,4e-169;
 Matches 445; Conservative 21; Mismatches 46; Indels 2; Gaps 2;
 QY 1203 MANNFSDMHLFRSMVLEMSVESQVYDEDPREFLNENSSAPVPHGESARRSSTLELTEL 1262
 DB 1 MNNFPGDWHLFRSTVLEMRNAESHVYEDPREFLSESSSGAPVHGEARRASINELPHTEL 60
 QY 1263 SSOTPTTLNFSFELMTLGLDEGAPRHSNLSMWSOTRRTPSSLNLSQSSSTIEISKLDK 1322
 DB 61 SSOTPTTLNFSFELMTLGLDEGAPRHSNLSMWSOTRRTPSSLNLSQSSSTIEISKLDK 120
 QY 1323 VQAEYRDAYREYIAQMSOLEGCGSTTIGRSSPHSTTYIGQSSGSGSTHSTLEDEGRKE 1382
 DB 121 VQAEYRDAYREYIAQMSOLEGCGSTTIGRSSPHSTTYIGQSSGSGSTHSTLEDEGRKD 180
 QY 1383 GELKQEDGRKSFLLMKRGVDYIDYSSGVSSTNEASPLDPTTEDEKSDQSGSKLLPKKSS 1442
 DB 181 SEPKPDGKRKSFLLMKRGVDYIDYSSGVSSTNEASPLDPTTEDEKSDQSGSKLLPKKSS 240
 QY 1443 RPSLFQTDKLKGGGRLYQKLPSEDESGTGRVQITPHCSKMITRTKRLAKQRECAPOE 1502
 DB 241 RSLFQTDKLKGGGRLYQKLPSEDESGTGRVQITPHCSKMITRTKRLAKQRECAPOE 299
 QY 1503 HSAEPIRTFKAEYISDALLDKSDSGVRSNNESSPMHSLHNEAADSOLEKANLTEL 1562
 DB 300 HSAEPIRTFKAEYISDALLDKSDSGVRSNNESSPMHSLHNEAADSOLEKANLTEL 359
 QY 1563 EDEGHSGRKMPHSLSLGLDPTIARMSICSEDKKSPSECSLIASSPEESWPAQORAYNLN 1622
 DB 360 EDDSHSGKRGIPIHSLSLGLDPTIARMSICSEDKKSPSECSLIASSPEESWPAQORAYNLN 419
 QY 1623 RTPSTYTLNNNTAPTRANQNPDEIGIETSYVILRPQSPRPTRAVQENELKNAHKRS 1682
 DB 420 RTPSTYTLNNNSAPARANQNPDEIGIETSYVILRPQSPRPTRAVQENELKNAHKRS 479
 QY 1683 QRSSYTRLSKDASELH-AASSESTGFEERESTL 1715
 DB 480 QRSSYTRLSKDPPELHAAASSESTGFEERESTL 513
 RESULT 11
 ID ABG08697
 XX ABG08697 standard; Protein; 624 AA.
 XX
 AC ABG08697;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #8688.
 XX
 KW Human: chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 PF 30-MAR-2001; 2001WO-US08631.
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 PA (HYSE-) HYSEQ INC.
 XX
 PI Dermanac RT, Liu C, Tang YT;
 XX
 DR MPI: 2001-639362/73.
 DR N-PSDB: AAS72884.

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity

XX Claim 20: SEQ ID NO 39056; 103bp; English.

CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_sequences.

XX Sequence 624 AA;

Query Match 24.3%; Score 2150; DB 22; Length 624;

Best Local Similarity 80.6%; Pred. No. 3e-160;

Matches 435; Conservative 31; Mismatches 66; Indels 8; Gaps 7;

QY 1183 KANINGRVLSQCINDELKEMANFGDMLFRSMVLEMSRVSQVYVDEPRLNENSSAP 1242

DB 86 EAININGRVLAQCINDELKEMANFGDMLFRSTVLEMAHSHVYVDEPRLNENSSGP 145

QY 1243 VPHGESARRSSHHELPTELSSQPTTLNLFSEELNLTGLDEAPRHSNLSMOSQRRRP 1302

DB 146 APHGEPRARRASHHELPTELSSQPTTLNLFSEELNLTGLDEAPRHSNLSMOSQRRRP 205

QY 1303 SLSSINQDSSIEISKTLTKVOAEYRDYREYIAOMSQLEGSGSTISGRSSPHSTYYI 1362

DB 206 SLSSINQDSSIEISKTLTKVOAEYRDYREYIAOMSQLEGSGSTISGRSSPHSTYYI 265

QY 1363 GQSSSGSIHSTLEQERKGBELKQEDGRKSFMLKRGDVIYSSSGVSTNEASPLDITE 1422

DB 266 GQSSSGSIHSTLEQERKGBELKQEDGRKSFMLKRGDVIYSSSGVSTNEASPLDITE 325

QY 1423 EDESDQSGSKLPGKSSSRPSPFQDILKGGGLRYQVLPEDEDESGRGRIQTHPCS 1482

DB 326 EDESDQSGSKLPGKSSSRPSPFQDILKGGGLRYQVLPEDEDESGRGRIQTHPCS 384

QY 1483 KMITRKLKAKOREKCAPSEHSAEPIRTFKAKEVYSLDALDKK-DSSDS-GVRSNESP 1540

DB 385 KDKDKRKPEKVERKVERKSPHSEVPIRTFKAKEVYSLDALDKKNGSSDSWSDPRESSP 444

QY 1541 NH-SLHNEAAD-SQLEKAMVLEDEGSHGKRGMPHSLSGLODPTIARMSICSEDKSP 1598

DB 445 NHVSAQCKCADDLPLDKRNLIELEDDSHGKRGIPHSGLDPTIARMSICSEDKSP 504

QY 1599 SEGLT--ASSPESWACAKATNLNRTPTSTVTLNNTATNANONFDEIBETISRTSOY 1656

DB 505 SRLASLDXPSPEENWACAKATNLNRTPTSTVTLNNTATNANONFDEIBETISRTSOY 564

QY 1657 ILRGPSPPLAVONENLKMMAHRSQSSYTRLSKASPLHAH-SESTGFPEERSTIL 1715

DB 565 ILRSPSPNPTTIONENLKMMAHRSQSSYTRLSKASPLHAH-SESTGFPEERSTIL 624

RESULT 12

AA040811

ID AA040811 standard; Protein; 120 AA.

XX AA040811;

DT 22-OCT-2001 (first entry)

DE Human polypeptide SEQ ID NO 5742.

XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;

XX peripheral nervous system; neuropathy; central nervous system; CNS;

XX Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;

XX amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;

XX chemokine; thrombolytic; drug screening; arthritis; inflammation;

XX leukaemia.

XX Homo sapiens.

XX WO200153312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000WO-US34263.

XX 21-JAN-2000; 2000US-0488725.

XX 25-APR-2000; 2000US-0552317.

XX 09-JUL-2000; 2000US-0598042.

XX 19-JUL-2000; 2000US-0620312.

XX 03-AUG-2000; 2000US-0653450.

XX 14-SEP-2000; 2000US-0662191.

XX 19-OCT-2000; 2000US-0693036.

XX 29-NOV-2000; 2000US-0727344.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;

XX Zhao QA, Zhou P, Goodrich R, Drmanac RT;

XX WPI; 2001-442253/47.

XX N-PSDB; AA159967.

XX Novel nucleic acids and polypeptides, useful for treating disorders

XX such as central nervous system injuries -

XX Example 2: SEQ ID NO 5742; 10078bp; English.

XX The invention relates to human nucleic acids (AA157798-AA161369) and

XX the encoded polypeptides (AA038642-AA042213) with nootropic,

XX immunosuppressant and cytostatic activity. The polynucleotides are useful

XX in gene therapy. A composition containing a polypeptide or polynucleotide

XX system, such as peripheral nervous injuries, peripheral neuropathy and

XX localised neuropathies and central nervous system diseases, such as

XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic

XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the

XX utilisation of the activities such as: immune system suppression,

XX activating/inhibiting activity, chemotactic/chemokinetic activity, haemostatic

XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,

XX assays for receptor activity, arthritis and inflammation, leukaemias and

XX C.N.S disorders.

XX Note: The sequence data for this patent did not form part of the printed

XX specification.

XX Sequence 120 AA;

XX Query Match 6.9%; Score 614; DB 22; Length 120;

XX Best Local Similarity 99.2%; Pred. No. 2.6e-40;

XX Matches 119; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX 217 TOSVKEILKRNPNVNLTKDGNWALMIAKGEHIEIVODLDAGTYVNIPIRSGPTVLIG 276

Db 1 TOSVKEILKRNPNVNLTKDGNFALMISKEGHTEIVODLDAGTYVNPDRSGTVLIG 60
QY 277 AARGHVEIVRALLQKYADIDIRGDNKNTALYMAVEKGNATVRIILOCNPDTEICTDG 336
61 AARGHVEIVRALLQKYADIDIRGDNKNTALYMAVEKGNATVRIILOCNPDTEICTDG 120

RESULT 13
AAB94785
ID AAB94785 standard; Protein: 129 AA.
AC AAB94785;
DT 26-JUN-2001 (first entry)
XX
DE Human protein sequence SEQ ID NO:15890.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
OS Homo sapiens.
XX
XX EPI074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
DR WPI: 2001-318749/34.
XX
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
PS Claim 8; SEQ ID 15890; 2537PP + CD ROM; English.
XX
XX
CC The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination
CC of the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
XX
SQ Sequence 129 AA;
Query Match 6.8%; Score 599.5; DB 22; Length 129;

Best Local Similarity 90.7%; Pred. No. 4.2e-39;
Matches 117; Conservative 4; Mismatches 7; Indels 1; Gaps 1;
QY 1588 MSICSEDKSPSECSLIASSPESWPACQAYNLNRTPTVTNNAPTNRANQFDEI 1647
DB 1 MSICSEDKSPSECSLIASSPESWPACQAYNLNRTPTVTNNAPTNRANQFDEI 60
QY 1648 EGIRETSQVILRPSPNPTFANQENIKSAHRRSQSSYTRISKASELH-AASESFG 1706
DB 61 EGIRETSQVILRPSPSPPTTTONENIKSWHRRSQSSYTRISKDPPELHAASSESTG 120
QY 1707 FGEERESIL 1715
DB 121 FGEERESIL 129

RESULT 14
AAB56277
ID AAB56277 standard; Protein: 109 AA.
AC AAB56277;
DT 13-MAR-2001 (first entry)
XX
XX
DE Human secreted protein sequence encoded by gene 82 SEQ ID NO:371.
XX
KW Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;
KW antiinflammatory; antiproliferative; cytostatic; cardiant; vasotropic;
KW cerebroprotective; neurotropic; neuroprotective; antibacterial; vitucide;
KW fungicide; ophthalmological; gene therapy; pathological condition;
KW autoimmune disease; rheumatoid arthritis; hyperproliferative disorder;
KW neoplasm; cardiovascular disorder; cardiac arrest; cerebral ischemia;
KW cerebrovascular disorder; angiogenesis; nervous system disorder;
KW Alzheimer's disease; infection; ocular disorder; corneal infection;
KW wound healing; skin aging; food additive; preservative.
XX
OS Homo sapiens.
XX
XX
PN WO200070042-A1.
XX
XX 23-NOV-2000.
XX
PF 11-MAY-2000; 2000WO-0512788.
XX
PR 13-MAY-1999; 99US-0134068.
XX
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM, Moore PA, Young PE, Komatsoulis GA, Birse CE;
PI Duan RD, Florence KA, Soppet DR;
DR WPI: 2000-679828/66.
XX
XX
PT Isolated nucleic acid molecule encoding a human secreted protein is
PT used in preventing, treating or ameliorating a medical condition -
XX
PS Disclosure; Page 1025-1026; 1065PP; English.
XX
XX
CC The polynucleotide sequences given in AAC99818 to AAC99977 encode the
CC human secreted proteins given in AAB56077 to AAB56362. Human secreted
CC proteins have activities based on the tissues and cells the genes are
CC expressed in. Examples of activities include: immunosuppressive;
CC antiarthritic; antiinflammatory; antiproliferative; cytostatic; cardiant;
CC vasotropic; cerebroprotective; neurotropic; neuroprotective; antibacterial;
CC vitucide; fungicide; and ophthalmological. The human secreted
CC polynucleotides and proteins can be used to prevent, treat or ameliorate
CC a medical condition in e.g. humans, mice, rabbits, goats, horses, cats,
CC dogs, chickens or sheep. They are also used in diagnosing a pathological
CC condition or susceptibility to a pathological condition. Disorders which
CC are diagnosed or treated include autoimmune diseases e.g. rheumatoid
CC arthritis, hyperproliferative disorders e.g. neoplasms of the breast or
CC liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular
CC disorders e.g. cerebral ischemia, angiogenesis, nervous system disorders

CC	e.g. Alzheimer's disease, infections caused by bacteria, viruses and	CC	302 DNKTALYMAVEKGNATMVADIIQCNPDEIETCKDSETPLIKATMKRNIEVEYELLDGAK	361
CC	fungi and ocular disorders e.g. corneal infection. The proteins can also	CC		CC
CC	be used to aid wound healing and epithelial cell proliferation, to	CC		CC
CC	prevent skin aging due to sunburn, to maintain organs before	CC		CC
CC	transplantation, for supporting cell culture of primary tissues, to	CC		CC
CC	regenerate tissues and in chemotaxis. The proteins can also be used as a	CC		CC
CC	food additive or preservative to increase or decrease storage	CC		CC
CC	capabilities. AAC99809 to AAC99817 and AAB5076 represent sequences used	CC		CC
CC	in the exemplification of the present invention.	CC		CC
XX		XX		XX
SQ	Sequence 109 AA:			
	Query Match 6.2%; Score 549; DB 21; Length 109;			
	Best Local Similarity 98.1%; Pred. No. 3e-35;			
	Matches 105; Conservative 2; Mismatches 0; Indels 0; Gaps 0			
QY				
b	302 DNKTALYMAVEKGNATMVADIIQCNPDEIETCKDSETPLIKATMKRNIEVEYELLDGAK	361		
	1 DNKTALYMAVEKGNATMVADIIQCNPDEIETCKDSETPLIKATMKRNIEVEYELLDGAK	60		
QY				
	362 VSAVDKKGDTPLHVAIRGSRRIAEELLRNPKDGRLLYRPNKAGETP	408		
	61 VSAVDKKGDTPLHVAIRGSRRIAEELLRNPKDGRLLYRPNKAGETP	107		
Db				
RESULT 15				
AAO20513				
XX	AAO20513 standard; Protein; 1724 AA.			
XX				
AC	AAO20513;			
XX				
DT	27-JUN-2002 (first entry)			
XX				
DE	Protein of APP related human homologue hCP51594.			
XX				
KW	Neuroprotective; nootropic; transgenic fly; Alzheimer's disease; Abeta;			
KW	amyloid precursor protein; tissue-specific expression control; human APP;			
XX	APP pathway modulator; gene therapy.			
XX				
OS	Homo sapiens.			
XX				
PN	WO200226820-A2.			
XX				
PD	04-APR-2002.			
XX				
PF	01-OCT-2001; 2001MO-EP11345.			
XX				
PR	29-SEP-2000; 2000US-236893P.			
XX	14-JUN-2001; 2001US-298309P.			
XX				
PA	(NOVS) NOVARTIS AG.			
PA	(NOVS) NOVARTIS-ERFINDUNGEN VERW GES MBH.			
XX				
PI	Cohen D, Dengler UJ, Finelli AL, Freuler F, Konsolaki M;			
PI	Reinhardt MMHM, Zusan S;			
XX				
DR	WPI: 2002-315796/35.			
XX	N-PSDB; AAK99407.			
XX				
PT	New transgenic fly, containing DNA encoding an Abeta portion of human			
PT	APP, useful for identifying agents which modulate the APP pathway and			
PT	which can be used to treat Alzheimer's disease -			
XX				
PS	Example 4; Page 116-119; 129pp; English.			
XX				
XX	The invention relates to a transgenic fly whose genome comprises DNA			
CC	encoding a polypeptide having the Abeta portion of human amyloid			
CC	precursor protein (APP), fused to a signal sequence. The DNA sequence			
CC	encodes a 123 (Abeta40) or 129 (Abeta42) amino acid sequence, given in			
CC	the specification. The DNA sequence is operably linked to a tissue-			
CC	specific expression control sequence. Expression of the sequence gives			
CC	the fly an altered phenotype. The purpose of the invention is for			
CC	identifying agents that inhibit or promote the expression and/or function			

Query Match	6.1%; Score 538.5; DB 23; Length 1724;
Best Local Similarity	20.3%; Pred. NO. 2.4e-32;
Matches 399;	Conservative 250; Mismatches 661; Indels 621; Gaps 722
QY 14	EEENIDALKALLKCKDYDERNECGTPRLIMAEQVNEIYVELLKGANCNLEDDNNT 73
Db 45	KEGHVEVSEELLQREANDAAATKRGNTALHTASLAGQAEVVKVLTNGAVNANSONGPT 104
QY 74	ALLSAKKEGHIIHVEELLKSGASLEHRMGVNTALMAACKRGIDVYELL 124
Db 105	PLYMAAOENHLEVVFKLLDNGASQSLATEDGFTPLAVALQOGHVOVVLLLENDTKGVR 164
QY 125	-----SHGANPSVTGL----- 135
Db 165	LPALHTAAKDKDTKAALLLLQNDNNADYESGFTPLHIAHNGINIVATLLNRAAVD 224
QY 136	---QYSVPELIWAAGRGADIVHLLDLQGAQVNCSDKXTGTPPLWAARKG----- 182
Db 225	FTARNIDITPLHVAASKRGANNAVYKLLLDGKAKIDAKTBDGILPLHGASRSHQEVEMLLD 284
QY 183	-----HLECYKHLIAMGADVDQGANSMATLIVAKGQGYTOS 219
Db 285	RAAPILSKTRKNGLSPLHMATQGDHNCYOLLQHNVPDVTYNDYLALHVAACHGNYV 344
QY 220	VKEILKRNPNVNLTDKQDGFALMIAKSGHEIYODLLDAGTYNIPDRSGDVLIGAVR 279
Db 345	AKVLLDKRANPNNAKALNGFTPLHIACKNRKIKVLELLKHGASIQAVTERGETALMAAR 404
QY 280	GGHVEIVRALLRKYADIDIRGQDNATLYUMVAVEGNATPVNADILQCNPDTEICTKGEPR 339
Db 405	SGQAEVYRVLYDQGAQVEAKAKDDQTPHIAIRLKGKADIDYQQLLDQASPMATTSGYTPR 464
QY 340	LKATMKRNIIEVELLDKGAKVASADKGDTPHVAIRGRSRRLAEILLRNPDKGRLLY 399
Db 465	LHTSAREGHEVDAAFLIDHGASLSTTTKSGTLPLHVAHNDQKVALTL---DQASPH 521
QY 400	RPKKAEFTYNDIC-SHOKSILTQI--FGAR-----HLSSTEDGMLGDYLX 444
Db 522	AAAKNGYTPLHIAAKKNQMDIATTLLEEGADANAVTROGIAVHIAQEGHVDMSYLLG 581
QY 445	SSALADILSEPTMOPPICVGLAQMGSGKSFLLKLEDEK-----TFAGQOTE-- 493
Db 582	RNANVNLNKSGLTP---LHLAAQ-----EDRVNVAEVLVNGCAHVDQOTKY 626
QY 494	-----PLFQPSWL--IVFTLLLCGGLGVFAFVDYMTALAIASL 531
Db 627	GRPLRPHGKECVHLVANGYTPPLHQAAQGGHIIINVLQNN-----ASRPELVVYTEKH 681
QY 532	SFLALITYIFVIYFGGRREGESNMWMAALSTRLARHIGYLELLFKIMEVNP----- 583
Db 682	KM-----NVPETNNEVLDMDEDEVRKANAPEMLSDEYISDVEBGNRCT 725
QY 584	-----PELPDQTKALPVRFPLTDVNRSSGEGESLA-----EMATLSD--ACE 627
Db 726	WYIIPVQVEFTYKTDTFKREAFDVGLLSTSGAEDMAIGDPTKYLPODLKELSDSLAE 785
QY 628	REFGF---LATRLFRVFETESQKKKKMK-----TCL- 658
Db 786	GYMGFSLGARSARPLVSEFVNDARGSMGSHHGMRIIPRRCSTAPTRTCTLVYRHL 845
QY 659	-----PSFVIFLTVGCIITAGITLLATRPVDPKHLTVNALLISTA 698
Db 846	ANPPVVEEGGLASRLVEKGPAGQFL---GPVIEIPLPHGSGMRKEKRELIV-----LR 896

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2003, 14:34:04 ; Search time 55.5 Seconds

(without alignments)
4117.564 Million cell updates/sec

Title: US-10-021-571-4

Perfect score: 8884

Sequence: 1 MSVLISQSVINVEENIPAA.....LHAASSSTGGEERESIL 1715

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A.Geneseq_101002.*
1: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
21: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8884	100.0	1715	22	AAAM38993
2	8884	100.0	1715	22	AAAM39025
3	8884	100.0	1715	23	AAU96840
4	8362.5	94.1	1762	23	AAU96841
5	8301	93.4	1763	23	AAU80244
6	3050.5	34.3	705	22	AAU75604
7	2851.5	32.1	551	22	AAU75604
8	2791	31.4	543	22	AAU95191
9	2638	29.7	513	22	AAU03645
10	2503	28.2	624	22	ABG08697

11	2379.5	26.8	1498	22	ABB64857
12	666	7.5	129	22	AAU94785
13	620	7.0	120	22	AAU0811
14	553	6.2	109	21	AAU56277
15	526	5.9	187	22	AAU79160
16	521	5.9	4274	22	ABG00972
17	519.5	5.8	2443	22	ABB60521
18	519.5	5.8	4386	22	ABG07375
19	519	5.8	1724	23	AAU0513
20	514	5.8	4397	22	ABG21944
21	499.5	5.6	187	22	AAU0779
22	482.5	5.4	784	22	ABB62206
23	480.5	5.4	1745	19	AAU70608
24	480.5	5.4	1745	19	AAU76776
25	480.5	5.4	1745	21	AAU11589
26	480.5	5.4	1745	23	AAU021368
27	478.5	5.4	1549	22	ABB58328
28	478.5	5.4	1549	22	ABB67412
29	463	5.2	2119	22	ABB64823
30	447	5.0	743	23	ABB97453
31	446	5.0	1133	22	ABG08071
32	440	5.0	378	22	ABB64963
33	438.5	4.9	1053	22	AAU28174
34	438	4.9	1088	22	AAU62429
35	438	4.9	1097	22	ABG21941
36	437	4.9	740	22	AAU01030
37	437	4.9	1569	23	AAU74360
38	430.5	4.8	705	22	AAU20496
39	429	4.8	435	22	AAU66710
40	429	4.8	435	22	AAU93879
41	429	4.8	627	23	AAU17136
42	421	4.7	1377	22	ABG08072
43	417.5	4.7	2441	22	ABB62231
44	415.5	4.7	1267	22	ABG22790
45	415	4.7	378	22	ABB64962

ALIGNMENTS

RESULT 1	
AAAM38993	
ID	AAAM38993 standard; Protein: 1715 AA.
XX	
AC	AAAM38993;
XX	
DT	22-OCT-2001 (first entry)
XX	
DE	Human polypeptide SEQ ID NO 2138.
XX	
KW	Human; nocotropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW	peripheral nervous system; neuropathy; central nervous system; CNS;
KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; Chemotactic;
KW	chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW	leukaemia.
XX	
OS	Homo sapiens.
XX	
PN	WO200153312-A1.
XX	
PD	26-JUL-2001.
XX	
PF	26-DEC-2000; 2000MO-US34263.
XX	
PR	21-JAN-2000; 2000US-0488725.
PR	25-APR-2000; 2000US-0552317.
PR	09-JUL-2000; 2000US-0598042.
PR	19-JUL-2000; 2000US-0620312.
PR	03-AUG-2000; 2000US-0653450.
PR	14-SEP-2000; 2000US-0662191.
PR	19-OCT-2000; 2000US-0693036.
PR	29-NOV-2000; 2000US-0727344.

QY 1621 NRPESTVTLNNSAPANRANONFDEMEGIRETSQVILRPSSPPPTTIONENLKSMTKRR 1680
 DB 1621 NRPESTVTLNNSAPANRANONFDEMEGIRETSQVILRPSSPPPTTIONENLKSMTKRR 1680
 QY 1681 SQRSSYRLSKDPPELHAASSESTGGEERESTL 1715
 DB 1681 SQRSSYRLSKDPPELHAASSESTGGEERESTL 1715

RESULT 2
 AAM39025
 ID AAM39025 standard; Protein; 1715 AA.
 AC AAM39025;
 XX 22-OCT-2001 (first entry)
 DT Human polypeptide SEQ ID NO 2170.
 DE Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW Peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
 KW Leukaemia.
 OS Homo sapiens.
 XX MO20015312-A1.
 PN 26-JUL-2001.
 PD 26-DEC-2000; 2000WO-US34263.
 PE 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.

(HYSE-) HYSEQ INC.
 FA Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX WPI: 2001-442253/47.
 DR N-PSDB: AA158181.
 PS Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 XX Example 4: SEQ ID NO 2170; 10078pp; English.
 PS The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AAM38642-AA42213) with nootropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.

XX Sequence 1715 AA.
 SQ Query Match 100.0%; Score 8884; DB 22; Length 1715;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSYLIQSQSYINVEENIRALAKLEKCDVDERNCCGTPIMIAEGCNLEIVELIKN 60
 DB 1 MSYLIQSQSYINVEENIRALAKLEKCDVDERNCCGTPIMIAEGCNLEIVELIKN 60
 QY GANCNLEDDNNMTALISASKEGHHVIEELIKGVNLERDGGMTALWMACTKGRTPVY 120
 DB GANCNLEDDNNMTALISASKEGHHVIEELIKGVNLERDGGMTALWMACTKGRTPVY 120
 QY 61 GANCNLEDDNNMTALISASKEGHHVIEELIKGVNLERDGGMTALWMACTKGRTPVY 120
 DB 61 GANCNLEDDNNMTALISASKEGHHVIEELIKGVNLERDGGMTALWMACTKGRTPVY 120
 QY 121 ELLISHGANDSVTGLQYSVYPIIWAAGRHADIVHLLQNGAKVNCSDKYGTTPVMAAR 180
 DB 121 ELLISHGANDSVTGLQYSVYPIIWAAGRHADIVHLLQNGAKVNCSDKYGTTPVMAAR 180
 QY 181 KGHLEGVKHLAMGADVDEGANSMTALIVAKGGTOSVKRILKRNPNVNLTDGNTA 240
 DB 181 KGHLEGVKHLAMGADVDEGANSMTALIVAKGGTOSVKRILKRNPNVNLTDGNTA 240
 QY 241 LMIAKRGHTEIVODLLDAGTYVNIIPDRSGDVIIGAVNGHVEIVRALLQKADIDIRG 300
 DB 241 LMIAKRGHTEIVODLLDAGTYVNIIPDRSGDVIIGAVNGHVEIVRALLQKADIDIRG 300
 QY 301 QDNKTALVNAVEKGNATVNRDILQCNPDTEICTKGEPLIKATKRNIEVELLDKGA 360
 DB 301 QDNKTALVNAVEKGNATVNRDILQCNPDTEICTKGEPLIKATKRNIEVELLDKGA 360
 QY 361 KSAVADKKGDTPLHAIIRGRSRKLAELLRNKDGLYRPPKAGETPNIDCSHOKSL 420
 DB 361 KSAVADKKGDTPLHAIIRGRSRKLAELLRNKDGLYRPPKAGETPNIDCSHOKSL 420
 QY 421 TQIFGARHLSPETDMDLGYDLYSSALADILSEPTMPPICVGYAOMGSGKSPFLKL 480
 DB 421 TQIFGARHLSPETDMDLGYDLYSSALADILSEPTMPPICVGYAOMGSGKSPFLKL 480
 QY 481 EDEMKTFAQOIPEPLFOFSWLVLETLGLCGGLGPAFTVHPNIGIAVSFLALYIF 540
 DB 481 EDEMKTFAQOIPEPLFOFSWLVLETLGLCGGLGPAFTVHPNIGIAVSFLALYIF 540
 QY 541 FIYIYFGRRGREGSNMAMVLSRLARHIGYELLLKMFVNPPELPEOTKALVRFLE 600
 DB 541 FIYIYFGRRGREGSNMAMVLSRLARHIGYELLLKMFVNPPELPEOTKALVRFLE 600
 QY 601 TDYNRLSVSGGETSLAEMITATLSDACEREFGLATRLRFVFTEDTQKKKKKTCCLPS 660
 DB 601 TDYNRLSVSGGETSLAEMITATLSDACEREFGLATRLRFVFTEDTQKKKKKTCCLPS 660
 QY 661 FVIFLEFICITISGTLIAIFRPVDPKHLVNAVVISASVGLAVLNCRTMVOYLDLIL 720
 DB 661 FVIFLEFICITISGTLIAIFRPVDPKHLVNAVVISASVGLAVLNCRTMVOYLDLIL 720
 QY 721 NSQKRHLHNAASKLHLKSEGFMYLKCCEVELAMAMAKTIDSFOTNORLVIYIDGLDCA 780
 DB 721 NSQKRHLHNAASKLHLKSEGFMYLKCCEVELAMAMAKTIDSFOTNORLVIYIDGLDCA 780
 QY 781 EODKVLQMDLTVRVLFSGKPTAIFASDPHIIKAIINONLSVLRDSNINGHYRNIVH 840
 DB 781 EODKVLQMDLTVRVLFSGKPTAIFASDPHIIKAIINONLSVLRDSNINGHYRNIVH 840
 QY 841 LVPFLNSRGLSNARKFLVTSANGVPCSDTGTGIEDDADRVRSQSLDEMKRLGSKTALN 900
 DB 841 LVPFLNSRGLSNARKFLVTSANGVPCSDTGTGIEDDADRVRSQSLDEMKRLGSKTALN 900
 QY 901 RRDYRRRQMTITRQMSFDLTKLVYEDWFSIDSPTMRRLNIVSVTRLLRANQIS 960
 DB 901 RRDYRRRQMTITRQMSFDLTKLVYEDWFSIDSPTMRRLNIVSVTRLLRANQIS 960
 QY 961 FWMDRLASINLITEOPRTSWLLIYLBETBSIPQOMTKTYERISNNIPTTKDVEPL 1020
 DB 961 FWMDRLASINLITEOPRTSWLLIYLBETBSIPQOMTKTYERISNNIPTTKDVEPL 1020

```

Db      961  FNMRLASWMLTEQWMPRTSMILLEYETEGIPDQMLTKTYERISKNIPTTNDVPELL 1020
QY      1021  EIDGDINFEVFLSSRIPLVAVARYKVFPLCTVNLDPKLEITADVAAAEQISIGLAY 1080
Db      1021  EIDBDIRNFEEVLSRIPVLARVAVKFLPCTVNLDPKLEITADVAAAEQISIGLAY 1080
QY      1081  PPLPLHEGPAPRAGSYSPSVCSSTSPNGFAGVAVSPHSSYSGMTGPPQPFNRG 1140
Db      1081  PPLPLHEGPAPRAGSYSPSVCSSTSPNGFAGVAVSPHSSYSGMTGPPQPFNRG 1140
QY      1141  SGAPGAVVLLNSLNDVAVCEKIQIGLDQSMPLPYCTTITKANINGRVLAQCNDIELK 1200
Db      1141  SGAPGAVVLLNSLNDVAVCEKIQIGLDQSMPLPYCTTITKANINGRVLAQCNDIELK 1200
QY      1201  KEMNNFGDWMLFSTVLEMNASHVYVPEDPRLSESSSGPAPHGEPARASINELPHT 1260
Db      1201  KEMNNFGDWMLFSTVLEMNASHVYVPEDPRLSESSSGPAPHGEPARASINELPHT 1260
QY      1261  ELSQTPYTLNFSFEELNTGLDGAAPRHSNLSMOSQTRTPSLSLNSQDSIEISKL 1320
Db      1261  ELSQTPYTLNFSFEELNTGLDGAAPRHSNLSMOSQTRTPSLSLNSQDSIEISKL 1320
QY      1321  DKVOAERYDAYREYIAQMSQLEGPGSTTIGRSSPHSTYYMGSSGSIHSLMLEQEK 1380
Db      1321  DKVOAERYDAYREYIAQMSQLEGPGSTTIGRSSPHSTYYMGSSGSIHSLMLEQEK 1380
QY      1381  KDSEKPPDGRKSPFLMKRGDIVIDSSGVSSTNDASPLDPTTEDEKSDQSGSKLLPKK 1440
Db      1381  KDSEKPPDGRKSPFLMKRGDIVIDSSGVSSTNDASPLDPTTEDEKSDQSGSKLLPKK 1440
QY      1441  SERSLFQTDKLKSGGLRYOKLPSEDESGTESDNTPLDKDKRAEKGVYRVPKSP 1500
Db      1441  SERSLFQTDKLKSGGLRYOKLPSEDESGTESDNTPLDKDKRAEKGVYRVPKSP 1500
QY      1501  EHSAPRTFTKAREYISDALLDKDSSDVSRSSESSPNLSLNEVADDSQLEKANLIE 1560
Db      1501  EHSAPRTFTKAREYISDALLDKDSSDVSRSSESSPNLSLNEVADDSQLEKANLIE 1560
QY      1561  LEDDSHSGKRGIPHSLSGLDPIIARMSICSEDKKSPSECSLIASSPEENPACOKAYNL 1620
Db      1561  LEDDSHSGKRGIPHSLSGLDPIIARMSICSEDKKSPSECSLIASSPEENPACOKAYNL 1620
QY      1621  NRTPTVTLLNNSAPANRANONFDEMEGIRETSQVILRPSSPPTTIONENLKSMTIKR 1680
Db      1621  NRTPTVTLLNNSAPANRANONFDEMEGIRETSQVILRPSSPPTTIONENLKSMTIKR 1680
QY      1681  SORSSYTRLKSDPPELHAASSESTGFEERESIL 1715
Db      1681  SORSSYTRLKSDPPELHAASSESTGFEERESIL 1715

RESULT 3
AAU96840 standard: Protein; 1715 AA.
XX      AAU96840;
XX
XX      30-JUL-2002 (first entry)
XX
XX      Human kidins220 protein.
XX
XX      kidins220; kinase D interacting substrate of 22kDa; cytosolic;
XX      neuroprotective; gene therapy; protein kinase D; PKD; cancer;
XX      neurodegenerative disease; glioblastoma multiforme; prostate cancer;
XX      human.
XX
XX      Homo sapiens.
XX
XX      WO200220786-A2.
XX
XX      14-MAR-2002.
XX
XX      06-SEP-2001; 2001WO-GB03977.

```

```

XX      06-SEP-2000; 2000US-230449P.
PR      (IMCR ) IMPERIAL CANCER RES TECHNOLOGY LTD.
PA      Schiavo G, Iglesias T;
XX      WPI: 2002-371879/40.
XX      N-Psdb; ABR51221.
XX
XX      Claim 33; Fig 11; 228pp; English.
XX
XX      The invention relates to polypeptide comprising a kinase D interacting
XX      substrate of 220 kDa (Kidins220) from rat and human or their
XX      fragment, variant or fusion provided that the protein is not the
XX      polypeptide encoded by Genbank Accession No. AB033076. Also included are
XX      a polynucleotide encoding kidins220 provided that polynucleotide is
XX      not any one of the clones corresponding to the 61 Genbank Accession Nos.
XX      given in the specification; an expression/replicable vector comprising
XX      the polynucleotide; a recombinant host cell containing the polynucleotide
XX      or vector; an anti-Kidins220 antibody (used in the preparation of
XX      Kidins220); an agent (A1) which modulates activity of Kidins220;
XX      protein kinase D (PKD) or interaction between PKD and Kidins220;
XX      a mutant animal transgenic for kidins220; the use of an agent capable of
XX      detecting the expression of Kidins220 gene in manufacture of a diagnostic
XX      reagent for diagnosing or prognostic cancer or for monitoring the
XX      progression of cancer in a patient; a compound comprising a moiety that
XX      selectively binds to Kidins220 or its variant and another moiety; and a
XX      nucleic acid molecule encoding the compound. The antibody is useful for
XX      modulating the activity of Kidins220, where the antibody prevents a
XX      region of Kidins220 interacting from another protein, or reduces the
XX      ability of Kidins220 to bind to adenosine triphosphate (ATP). The
XX      Kidins220 polynucleotide is useful for identifying an agent which
XX      modulates the promoter activity of the polynucleotide, and Kidins220 is
XX      useful for identifying an agent which modulates activity of Kidins220,
XX      protein kinase D (PKD) and the interaction between PKD with Kidins220.
XX      Kidins220 or its polynucleotide is useful in medicine, e.g. using gene
XX      therapy or for diagnosing or monitoring progression of cancer in a
XX      patient. The agent is useful in manufacture of medicament, for use in
XX      treatment of neurodegenerative disease, and the agent which inhibits
XX      function of Kidins220 gene or its product is useful for treating cancer
XX      which is glioblastoma multiforme or prostate cancer. The compound is
XX      useful for imaging cancer in an individual, and for diagnosing or
XX      prognostic, and also for treating an individual. The present
XX      sequence represents Human Kidins220.
XX
XX      Sequence 1715 AA:
XX
XX      Query Match 100.0%; Score 8884; DB 23; Length 1715;
XX      Best Local Similarity 100.0%; Pred. No. 0;
XX      Matches 1715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

|||||
Db 241 LMIASKEGHEIYQDLDLAGTYVNIIPDRSGDFVLICAVGVGHEIYRALLQYADIDING 300
QY 301 QDNKTLAYAVENKGNATWTRDILQCPDPEICTKGEFTPLKATKRNIEVEVLLDKCA 360
Db 301 QDNKTLAYAVENKGNATWTRDILQCPDPEICTKGEFTPLKATKRNIEVEVLLDKCA 360
QY 361 KVSADVKKGDPTLHAIIRGRSRKLAELLRNPKDGLYRPKNACGTPPINIDCSHOKSL 420
Db 361 KVSADVKKGDPTLHAIIRGRSRKLAELLRNPKDGLYRPKNACGTPPINIDCSHOKSL 420
QY 421 TQIFGARHLSPETEDDMLGIDLYSSALADILSEPTMOPPICVGIYAQWGSKSFLLKTL 480
Db 421 TQIFGARHLSPETEDDMLGIDLYSSALADILSEPTMOPPICVGIYAQWGSKSFLLKTL 480
QY 481 EDEMKTFAQOQTEPLQFESWLVFLLLCGGIGLFAFTVHPNIGIANVSFLALTYE 540
Db 481 EDEMKTFAQOQTEPLQFESWLVFLLLCGGIGLFAFTVHPNIGIANVSFLALTYE 540
QY 541 FIYIFGRREGESNMAMAVLSTRLARHIGYLELLKLMFVNPEPEQTALPYRFLF 600
Db 541 FIYIFGRREGESNMAMAVLSTRLARHIGYLELLKLMFVNPEPEQTALPYRFLF 600
QY 601 TDYNRLSVGGETSLAEMTATLSDACERFGLATRLFRVETEDTQKKKKKTCCLPS 660
Db 601 TDYNRLSVGGETSLAEMTATLSDACERFGLATRLFRVETEDTQKKKKKTCCLPS 660
QY 661 FVLEFLIGIISGITLLAIFRVDPKHLVNAVLLSIAVGLAPVLCNRTWQVYLDL 720
Db 661 FVLEFLIGIISGITLLAIFRVDPKHLVNAVLLSIAVGLAPVLCNRTWQVYLDL 720
QY 721 NSQKRLHNAASKLHKLKSEGFMKVLCVEELMARAKTIDSEFTQNOTGLVYLIDGLDAC 780
Db 721 NSQKRLHNAASKLHKLKSEGFMKVLCVEELMARAKTIDSEFTQNOTGLVYLIDGLDAC 780
QY 781 EDDKVLQMDLYRVLFESKGPFLAIFASDPHHIIKAINQNLNSVLRDSNINHDMRNIVH 840
Db 781 EDDKVLQMDLYRVLFESKGPFLAIFASDPHHIIKAINQNLNSVLRDSNINHDMRNIVH 840
QY 841 LPEFLNSRGLSNARKFLVTSATNGDVPSCDTTGIOEDADRYSONSLGEMTKLGSSTALN 900
Db 841 LPEFLNSRGLSNARKFLVTSATNGDVPSCDTTGIOEDADRYSONSLGEMTKLGSSTALN 900
QY 901 RRDYRRROMQRTITRQMSFDLTKLVTEDWFSDSIPQMRRLNIVSYTGRLLRANQIS 960
Db 901 RRDYRRROMQRTITRQMSFDLTKLVTEDWFSDSIPQMRRLNIVSYTGRLLRANQIS 960
QY 961 FNMDELASWINLTEOMPRTSWLILEETEGIPDQMLKTIYERISKNIPTTKDVEPLL 1020
Db 961 FNMDELASWINLTEOMPRTSWLILEETEGIPDQMLKTIYERISKNIPTTKDVEPLL 1020
QY 1021 EIDGDIRNEFVFLSSTPVLAVADYVFLPCYVNDLPKREITIAVRAAREIOISIGLAY 1080
Db 1021 EIDGDIRNEFVFLSSTPVLAVADYVFLPCYVNDLPKREITIAVRAAREIOISIGLAY 1080
QY 1081 PPLPLEHGGPRAPSGTSPVSCSTSFNGPAGVVSPOPHSSYSGMTGQHPFYNG 1140
Db 1081 PPLPLEHGGPRAPSGTSPVSCSTSFNGPAGVVSPOPHSSYSGMTGQHPFYNG 1140
QY 1141 SCGAPGPVYLLNSLNDVACEKIKQIEGLDQSMLOYCTTIKANINGVLAQCNIDEIK 1200
Db 1141 SCGAPGPVYLLNSLNDVACEKIKQIEGLDQSMLOYCTTIKANINGVLAQCNIDEIK 1200
QY 1201 KEMNMPGDMHLEFRSTVLEMRNAESHVPEDEPRFSESSSGAPHEPPARRASHNLEPPT 1260
Db 1201 KEMNMPGDMHLEFRSTVLEMRNAESHVPEDEPRFSESSSGAPHEPPARRASHNLEPPT 1260
QY 1261 ELSSQPTPLNLSFEELNLTGLDEGAPRHSNLSWOSQTRRTPLSLNSQDSIEISKTL 1320
Db 1261 ELSSQPTPLNLSFEELNLTGLDEGAPRHSNLSWOSQTRRTPLSLNSQDSIEISKTL 1320
QY 1321 DKVOAETRDAYRYTIAQMSOLBEGPGSTTISGRSSHSTYWMQSSSGSISHSNLEQEK 1380
Db 1321 DKVOAETRDAYRYTIAQMSOLBEGPGSTTISGRSSHSTYWMQSSSGSISHSNLEQEK 1380

```

```

Db 1321 DKVOAETRDAYRYTIAQMSOLEGPGSTTISGRSSHSTYWMQSSSGSISHSNLEQEK 1380
QY 1381 KDSEPKPDGGRKSPFLMKRGDVIYSSGVSSTNDASPLDPTREDEKSDQSGSKLLPGKKS 1440
Db 1381 KDSEPKPDGGRKSPFLMKRGDVIYSSGVSSTNDASPLDPTREDEKSDQSGSKLLPGKKS 1440
QY 1441 SERSLFQTDKLKSGSLRYOKLPDEDESGTESDNTPLDKDKRAEKYERVPKSP 1500
Db 1441 SERSLFQTDKLKSGSLRYOKLPDEDESGTESDNTPLDKDKRAEKYERVPKSP 1500
QY 1501 EHSAPERTFTAKAYLSLDALLDKDSDSCVSRSESSPNHSLNEVAADSOLEKANLIE 1560
Db 1501 EHSAPERTFTAKAYLSLDALLDKDSDSCVSRSESSPNHSLNEVAADSOLEKANLIE 1560
QY 1561 LEDDSHSGKRGIPHSLSGLDPTIARMSICSEDKSPSECLIASPEENPAQOKAYNL 1620
Db 1561 LEDDSHSGKRGIPHSLSGLDPTIARMSICSEDKSPSECLIASPEENPAQOKAYNL 1620
QY 1621 NRTPTVTLLNNNSAPANRANONFDEMEGIRFTSOVILRPSSSPNPTTIONENLAKSMHNR 1680
Db 1621 NRTPTVTLLNNNSAPANRANONFDEMEGIRFTSOVILRPSSSPNPTTIONENLAKSMHNR 1680
QY 1681 SORSSYTRLSDPPELHAAASESTGCEERESIL 1715
Db 1681 SORSSYTRLSDPPELHAAASESTGCEERESIL 1715

RESULT 4
AA096841
ID AA096841 standard; Protein: 1762 AA.
XX
AC AA096841;
XX
DE 30-JUL-2002 (first entry)
XX
DE Rat kidneys220 protein from Genbank AA035185.
XX
KW Kidins220; kinase D interacting substrate of 22kDa; cytosolic;
KW neuroprotective; gene therapy; protein kinase D; PKD; cancer; rat;
KW neurodegenerative disease; glioblastoma multiforme; prostate cancer.
XX
OS Rattus sp.
XX
PN WO200220786-A2.
XX
PD 14-MAR-2002.
XX
PF 06-SEP-2001; 2001WO-GB03977.
XX
PR 06-SEP-2000; 2000US-230449P.
XX
PA (IMCR ) IMPERIAL CANCER RES TECHNOLOGY LTD.
XX
PI Schlavo G, Iglesias T;
XX
DR WPI: 2002-371879/40.
XX
PT Polypeptide kinase D interacting substrate of 220 kDa and
PT polynucleotide, useful for identifying modulators useful in treating
PT cancer and neurodegenerative diseases
XX
PS Disclosure; Fig 18; 228pp; English.
XX

The invention relates to polypeptide comprising a kinase D interacting
substrate of 220 kDa (Kidins220) from rat and human or their
fragment, variant or fusion provided that the protein is not the
polypeptide encoded by Genbank Accession No. AB033076. Also included are
a polynucleotide encoding Kidins220 provided that polynucleotide is
not any one of the clones corresponding to the 61 Genbank Accession Nos.
given in the specification; an expression/replicable vector comprising
the polynucleotide; a recombinant host cell containing the polynucleotide
or vector; an anti-Kidins220 antibody (used in the preparation of
Kidins220); an agent (AI) which modulates activity of Kidins220,

```

CC protein kinase D (PKD) or interaction between PKD and Kidins220:
 CC a mutant animal transgenic for Kidins220; the use of an agent capable of
 CC detecting the expression of Kidins220 gene in manufacture of a diagnostic
 CC reagent for diagnosing or prognosing cancer or for monitoring the
 CC progression of cancer in a patient; a compound comprising a moiety that
 CC selectively binds to Kidins220 or its variant and another moiety; and a
 CC nucleic acid molecule encoding the compound. The antibody is useful for
 CC modulating the activity of Kidins220, where the antibody prevents a
 CC region of Kidins220 interacting from another protein, or reduces the
 CC ability of Kidins220 to bind to adenosine triphosphate (ATP). The
 CC Kidins220 polynucleotide is useful for identifying an agent which
 CC modulates the promoter activity of the polynucleotide, and Kidins220 is
 CC useful for identifying an agent which modulates activity of Kidins220.
 CC protein kinase D (PKD) and the interaction between PKD with Kidins220.
 CC Kidins220 or its polynucleotide is useful in medicine, e.g. using gene
 CC therapy or for diagnosing or monitoring progression of cancer in a
 CC patient. The agent is useful in manufacture of medicament, for use in
 CC treatment of neurodegenerative disease, and the agent which inhibits
 CC function of Kidins220 gene or its product is useful for treating cancer
 CC which is glioblastoma multiforme or prostate cancer. The compound is
 CC useful for imaging cancer in an individual, and for diagnosing or
 CC prognosing, and also for treating an individual. The present
 CC sequence represents Rat Kidins220 protein from Genbank accession number
 CC AAG35185.

XX Sequence 1762 AA;

Query Match 94.1%; Score 8362.5; DB 23; Length 1762;

Best Local Similarity 91.6%; Pred. No. 0;

Matches 1616; Conservative 47; Mismatches 50; Indels 51; Gaps 4;

QY 1 MSVLISQSYINVEEENIPALKALEKCKVDNERNCGGTPMLIAEOGNLEIYKELKN 60
 DB 1 MSVLISQSYINVEEENIPALKALEKCKVDNERNCGGTPMLIAEOGNLEIYKELKN 60
 QY 61 GANCLIEDLDNNTALISASKEGHVHVEELKCGVNLERDMGNGTALMACYKGRDYY 120
 DB 61 GANCLIEDLDNNTALISASKEGHVHVEELKCGVNLERDMGNGTALMACYKGRDYY 120
 QY 121 ELLSHGANPSVTGLQSYVPIIWAAGRHADIVHLLONGKAVNCSKYGTPVYMAR 180
 DB 121 ELLSHGANPSVTGLQSYVPIIWAAGRHADIVHLLONGKAVNCSKYGTPVYMAR 180
 QY 121 ELLSHGANPSVTGLQSYVPIIWAAGRHADIVHLLONGKAVNCSKYGTPVYMAR 180
 DB 121 ELLSHGANPSVTGLQSYVPIIWAAGRHADIVHLLONGKAVNCSKYGTPVYMAR 180
 QY 181 KSHLECYKHLAMGADVDEGANSMATLVAAGGYTSYKELIKRNPNVNLTDKDGNTA 240
 DB 181 KSHLECYKHLAMGADVDEGANSMATLVAAGGYTSYKELIKRNPNVNLTDKDGNTA 240
 QY 241 LMIASKEGHEIVODLLAGTYVNPDRSGDTVLIGAVRGHVEIVRALLKYADIDIRG 300
 DB 241 LMIASKEGHEIVODLLAGTYVNPDRSGDTVLIGAVRGHVEIVRALLKYADIDIRG 300
 QY 240 LMIASKEGHEIVODLLAGTYVNPDRSGDTVLIGAVRGHVEIVRALLKYADIDIRG 299
 DB 240 LMIASKEGHEIVODLLAGTYVNPDRSGDTVLIGAVRGHVEIVRALLKYADIDIRG 299
 QY 301 QDNKPLVAWEKGNATWVRILOCNPTETICTKDGEPPLIKATMRNIEVEYELLDDGA 360
 DB 301 QDNKPLVAWEKGNATWVRILOCNPTETICTKDGEPPLIKATMRNIEVEYELLDDGA 360
 QY 300 QDNKPLVAWEKGNATWVRILOCNPTETICTKDGEPPLIKATMRNIEVEYELLDDGA 359
 DB 300 QDNKPLVAWEKGNATWVRILOCNPTETICTKDGEPPLIKATMRNIEVEYELLDDGA 359
 QY 361 KVSANDKGDTPHLAIIRGSRRLAELLRNPKDRLRYRPNKAGETPYNIDCSHQSKIL 420
 DB 361 KVSANDKGDTPHLAIIRGSRRLAELLRNPKDRLRYRPNKAGETPYNIDCSHQSKIL 420
 QY 360 KVSANDKGDTPHLAIIRGSRRLAELLRNPKDRLRYRPNKAGETPYNIDCSHQSKIL 419
 DB 360 KVSANDKGDTPHLAIIRGSRRLAELLRNPKDRLRYRPNKAGETPYNIDCSHQSKIL 419
 QY 421 TQIFGARHLSPTETGDMGLGYDLYSSALADILSEPTMQPPICVGLIYAOMGSKSKLLKL 480
 DB 421 TQIFGARHLSPTETGDMGLGYDLYSSALADILSEPTMQPPICVGLIYAOMGSKSKLLKL 480
 QY 420 TQIFGARHLSPTETGDMGLGYDLYSSALADILSEPTMQPPICVGLIYAOMGSKSKLLKL 479
 DB 420 TQIFGARHLSPTETGDMGLGYDLYSSALADILSEPTMQPPICVGLIYAOMGSKSKLLKL 479
 QY 481 EDEMKTFAGQOIEPLFOFSWILVFTLLCGGLGLFAFTVHPNIGVLSFLALYIF 540
 DB 481 EDEMKTFAGQOIEPLFOFSWILVFTLLCGGLGLFAFTVHPNIGVLSFLALYIF 540
 QY 480 EDEMKTFAGQOIEPLFOFSWILVFTLLCGGLGLFAFTVHPNIGVLSFLALYIF 539
 DB 480 EDEMKTFAGQOIEPLFOFSWILVFTLLCGGLGLFAFTVHPNIGVLSFLALYIF 539
 QY 541 FLVIYFGRRGESNMWAVLSTRLARHIGYELLLKLMFVNPPELPQTTKALPVRLE 600
 DB 541 FLVIYFGRRGESNMWAVLSTRLARHIGYELLLKLMFVNPPELPQTTKALPVRLE 600
 QY 540 FLVIYFGRRGESNMWAVLSTRLARHIGYELLLKLMFVNPPELPQTTKALPVRLE 599
 DB 540 FLVIYFGRRGESNMWAVLSTRLARHIGYELLLKLMFVNPPELPQTTKALPVRLE 599
 QY 601 TTYNRLSSVGETSLAEMATISDCERFGLARLRFRVKTETDQKKKKKTKCCPLS 660
 DB 601 TTYNRLSSVGETSLAEMATISDCERFGLARLRFRVKTETDQKKKKKTKCCPLS 660

DB 600 TTYNRLSSVGETSLAEMATISDCERFGLARLRFRVKTETDQKKKKKTKCCPLS 659
 QY 661 FVIFLEITIGITIGITILLAFRVPDKHLTVNAVILISIASVGLAFVLCNRTMQVILDSL 720
 DB 661 FVIFLEITIGITIGITILLAFRVPDKHLTVNAVILISIASVGLAFVLCNRTMQVILDSL 720
 QY 660 FVIFLEITIGITIGITILLAFRVPDKHLTVNAVILISIASVGLAFVLCNRTMQVILDSL 719
 DB 660 FVIFLEITIGITIGITILLAFRVPDKHLTVNAVILISIASVGLAFVLCNRTMQVILDSL 719
 QY 721 NSQKRRLHNASAKTHKLSGKFMKVLKCEVELMARNAKTIDSTQONTQRLVYIIDLDDG 780
 DB 721 NSQKRRLHNASAKTHKLSGKFMKVLKCEVELMARNAKTIDSTQONTQRLVYIIDLDDG 780
 QY 720 NSQKRRLHNASAKTHKLSGKFMKVLKCEVELMARNAKTIDSTQONTQRLVYIIDLDDG 779
 DB 720 NSQKRRLHNASAKTHKLSGKFMKVLKCEVELMARNAKTIDSTQONTQRLVYIIDLDDG 779
 QY 781 EODKVLMDLTVRVLFSGKGFIAIFASDPHIIIIKAIQNLNSVLRDSDNGHDMYRNIVH 840
 DB 781 EODKVLMDLTVRVLFSGKGFIAIFASDPHIIIIKAIQNLNSVLRDSDNGHDMYRNIVH 840
 QY 780 EODKVLMDLTVRVLFSGKGFIAIFASDPHIIIIKAIQNLNSVLRDSDNGHDMYRNIVH 839
 DB 780 EODKVLMDLTVRVLFSGKGFIAIFASDPHIIIIKAIQNLNSVLRDSDNGHDMYRNIVH 839
 QY 841 LPVFLNRGSLNARKFLVTSATNGDVCSPTTGIOEDAPRVQNSLGEFTKLGSTALN 900
 DB 841 LPVFLNRGSLNARKFLVTSATNGDVCSPTTGIOEDAPRVQNSLGEFTKLGSTALN 900
 QY 840 LPVFLNRGSLNARKFLVTSATNGDVCSPTTGIOEDAPRVQNSLGEFTKLGSTALN 899
 DB 840 LPVFLNRGSLNARKFLVTSATNGDVCSPTTGIOEDAPRVQNSLGEFTKLGSTALN 899
 QY 901 RRDYRRRQOMRTTRQMSFDLTKLVTEDMFSDISPQTRRLNIVSYTGRLLRANOIS 960
 DB 901 RRDYRRRQOMRTTRQMSFDLTKLVTEDMFSDISPQTRRLNIVSYTGRLLRANOIS 960
 QY 900 RRDYRRRQOMRTTRQMSFDLTKLVTEDMFSDISPQTRRLNIVSYTGRLLRANOIS 959
 DB 900 RRDYRRRQOMRTTRQMSFDLTKLVTEDMFSDISPQTRRLNIVSYTGRLLRANOIS 959
 QY 961 FMDRLASWNLLEQWPRYSWLLLYLEETEGIPDQWTLKTYERISKNIPTTKDYEPPL 1020
 DB 961 FMDRLASWNLLEQWPRYSWLLLYLEETEGIPDQWTLKTYERISKNIPTTKDYEPPL 1020
 QY 960 FMDRLASWNLLEQWPRYSWLLLYLEETEGIPDQWTLKTYERISKNIPTTKDYEPPL 1019
 DB 960 FMDRLASWNLLEQWPRYSWLLLYLEETEGIPDQWTLKTYERISKNIPTTKDYEPPL 1019
 QY 1021 EIDGDIINFVFLSSRTFVYARVYKFLCTYVNLDPKLEITIAVRAAREQISIGLAV 1080
 DB 1021 EIDGDIINFVFLSSRTFVYARVYKFLCTYVNLDPKLEITIAVRAAREQISIGLAV 1080
 QY 1020 EIDGDIINFVFLSSRTFVYARVYKFLCTYVNLDPKLEITIAVRAAREQISIGLAV 1079
 DB 1020 EIDGDIINFVFLSSRTFVYARVYKFLCTYVNLDPKLEITIAVRAAREQISIGLAV 1079
 QY 1081 PPLPLHGGPRAPSGYQOPSPVCSSTFNFPGFAGVVPSPHSSYSGMTGPOHPYRNR 1139
 DB 1081 PPLPLHGGPRAPSGYQOPSPVCSSTFNFPGFAGVVPSPHSSYSGMTGPOHPYRNR 1139
 QY 1080 PPLPLHGGPRAPSGYQOPSPVCSSTFNFPGFAGVVPSPHSSYSGMTGPOHPYRNR 1139
 DB 1080 PPLPLHGGPRAPSGYQOPSPVCSSTFNFPGFAGVVPSPHSSYSGMTGPOHPYRNR 1139
 QY 1140 -----GSG-----PAGGVYLL 1151
 DB 1140 -----GSG-----PAGGVYLL 1151
 QY 1140 FFAEYLYTPRYPGSOHLISRSSVKTSLPRDQNGPLPCSGGNNKROAVALPATGSLLL 1199
 DB 1140 FFAEYLYTPRYPGSOHLISRSSVKTSLPRDQNGPLPCSGGNNKROAVALPATGSLLL 1199
 QY 1152 NSLNVDAVCEKRLQIEGLDQSMPLPOYCTTKKANINGRYLAOCNIDELKEMAMNGDMH 1211
 DB 1152 NSLNVDAVCEKRLQIEGLDQSMPLPOYCTTKKANINGRYLAOCNIDELKEMAMNGDMH 1211
 QY 1200 SSMATVDVYCEKRLQIEGLDQSMPLPOYCTTKKANINGRYLAOCNIDELKEMAMNGDMH 1259
 DB 1200 SSMATVDVYCEKRLQIEGLDQSMPLPOYCTTKKANINGRYLAOCNIDELKEMAMNGDMH 1259
 QY 1212 LFRSTVLEMAHNAESHVVEDPRLSESSGAPRAGEPARASHNELPHTELSCQTPYTLN 1271
 DB 1212 LFRSTVLEMAHNAESHVVEDPRLSESSGAPRAGEPARASHNELPHTELSCQTPYTLN 1271
 QY 1260 LFRSMVLEMAHNAESHVVEDPRLSESSGAPRAGEPARASHNELPHTELSCQTPYTLN 1319
 DB 1260 LFRSMVLEMAHNAESHVVEDPRLSESSGAPRAGEPARASHNELPHTELSCQTPYTLN 1319
 QY 1272 FSPFELMTGLIDEGAPRSHNLSMOSQTRRTPSLSSLSNDSQSTIEISKLTDKVAEYRDY 1331
 DB 1272 FSPFELMTGLIDEGAPRSHNLSMOSQTRRTPSLSSLSNDSQSTIEISKLTDKVAEYRDY 1331
 QY 1320 FSPFELMTGLIDEGAPRSHNLSMOSQTRRTPSLSSLSNDSQSTIEISKLTDKVAEYRDY 1379
 DB 1320 FSPFELMTGLIDEGAPRSHNLSMOSQTRRTPSLSSLSNDSQSTIEISKLTDKVAEYRDY 1379
 QY 1332 REYIAQMSOLEGGGSGTIGSRSSPHSTYVMGSSGGSIHSMLEQEKDSEPKDDGR 1391
 DB 1332 REYIAQMSOLEGGGSGTIGSRSSPHSTYVMGSSGGSIHSMLEQEKDSEPKDDGR 1391
 QY 1380 REYIAQMSOLEGGGSGTIGSRSSPHSTYVMGSSGGSIHSMLEQEKDSEPKDDGR 1439
 DB 1380 REYIAQMSOLEGGGSGTIGSRSSPHSTYVMGSSGGSIHSMLEQEKDSEPKDDGR 1439
 QY 1392 KSFIMKRGVDYIDSSGVSSTNEASPLDPIIEDEEKSDQSSSKLLPKKSEERSLQOTDL 1499
 DB 1392 KSFIMKRGVDYIDSSGVSSTNEASPLDPIIEDEEKSDQSSSKLLPKKSEERSLQOTDL 1499
 QY 1440 KSFIMKRGVDYIDSSGVSSTNEASPLDPIIEDEEKSDQSSSKLLPKKSEERSLQOTDL 1511
 DB 1440 KSFIMKRGVDYIDSSGVSSTNEASPLDPIIEDEEKSDQSSSKLLPKKSEERSLQOTDL 1511
 QY 1452 KTKSGGLRYOKLPSEDESEGTTEESDNPPLKDKDKABEKAARVYKSPHSAEPIRTPI 1559
 DB 1452 KTKSGGLRYOKLPSEDESEGTTEESDNPPLKDKDKABEKAARVYKSPHSAEPIRTPI 1559
 QY 1500 KTKSGGLRYOKLPSEDESEGTTEESDNPPLKDKDKABEKAARVYKSPHSAEPIRTPI 1571
 DB 1500 KTKSGGLRYOKLPSEDESEGTTEESDNPPLKDKDKABEKAARVYKSPHSAEPIRTPI 1571
 QY 1512 KAKELYLDALDKKSDSGVRSSESSPNHSLNFEVADSOLEKANLIELEDDSHSGKGG 1571
 DB 1512 KAKELYLDALDKKSDSGVRSSESSPNHSLNFEVADSOLEKANLIELEDDSHSGKGG 1571
 QY 1560 KAKELYLDALDKKSDSGVRSSESSPNHSLNFEVADSOLEKANLIELEDDSHSGKGG 1619
 DB 1560 KAKELYLDALDKKSDSGVRSSESSPNHSLNFEVADSOLEKANLIELEDDSHSGKGG 1619
 QY 1572 IPHLSGLQDPIIARMSICSEDKKSPSEGLASPEENMPACOKAYNINRFPSTVTLNN 1631
 DB 1572 IPHLSGLQDPIIARMSICSEDKKSPSEGLASPEENMPACOKAYNINRFPSTVTLNN 1631
 QY 1620 MPHSLGLODPIIARMSICSEDKKSPSEGLASPEENMPACOKAYNINRFPSTVTLNN 1679
 DB 1620 MPHSLGLODPIIARMSICSEDKKSPSEGLASPEENMPACOKAYNINRFPSTVTLNN 1679
 QY 1632 NSAPANRANONPEDEGIRETSOYIIRPSSPMPPTIIONNLSKSMHKSORSRSTRLSK 1691
 DB 1632 NSAPANRANONPEDEGIRETSOYIIRPSSPMPPTIIONNLSKSMHKSORSRSTRLSK 1691
 QY 1680 NTAFTNANONPEDEGIRETSOYIIRPSSPMPPTIIONNLSKSMHKSORSRSTRLSK 1739
 DB 1680 NTAFTNANONPEDEGIRETSOYIIRPSSPMPPTIIONNLSKSMHKSORSRSTRLSK 1739

QY 1692 DPELHAASSTGFEERESTL 1715
 ID 1740 DASELH-AASSETGFEERESTL 1762

RESULT 5
 AA080244
 ID AA080244 standard; Protein: 1763 AA.

AC AA080244;

DT 30-JUL-2002 (first entry)

DE Rat Kidins220 protein.

Kidins220; kinase D interacting substrate of 22kDa; cytostatic;
 neuroprotective; gene therapy; protein kinase D; PKD; cancer; rat;
 neurodegenerative disease; glioblastoma multiforme; prostate cancer.

OS Rattus sp.

PN W020020786-A2.

PD 14-MAR-2002.

PF 06-SEP-2001; 2001WO-GB03977.

PR 06-SEP-2000; 2000US-230449P.

PA (IMCR) IMPERIAL CANCER RES TECHNOLOGY LTD.

PI Schiavo G, Iglesias T;

DR WPI: 2002-371879/40.

DR N-PSDB: ABK51214.

PT Polypeptide kinase D interacting substrate of 220 kDa and
 polynucleotide, useful for identifying modulators useful in treating
 cancer and neurodegenerative diseases

PS Claim 33; Fig 1; 228pp; English.

The invention relates to polypeptide comprising a kinase D interacting
 substrate of 220 kDa (Kidins220) from rat and human or their
 fragment, variant or fusion provided that the protein is not the
 polypeptide encoded by GenBank Accession No. AB033076. Also included are
 a polynucleotide encoding Kidins220 provided that polynucleotide is
 not any one of the clones corresponding to the 61 GenBank Accession Nos.
 given in the specification; an expression/replacable vector comprising
 the polynucleotide; a recombinant host cell containing the polynucleotide
 or vector; an anti-Kidins220 antibody (used in the preparation of
 Kidins220); an agent (AI) which modulates activity of Kidins220;
 protein kinase D (PKD) or interaction between PKD and Kidins220;
 a mutant animal transgenic for Kidins220; the use of an agent capable of
 detecting the expression of Kidins220 gene in manufacture of a diagnostic
 reagent for diagnosing or prognosing cancer or for monitoring the
 progression of cancer in a patient; a compound comprising a moiety that
 selectively binds to Kidins220 or its variant and another moiety; and a
 nucleic acid molecule encoding the compound. The antibody prevents a
 modulating the activity of Kidins220, where the antibody prevents a
 region of Kidins220 interacting from another protein, or reduces the
 ability of Kidins220 to bind to adenosine triphosphate (ATP). The
 Kidins220 polynucleotide is useful for identifying an agent which
 modulates the promoter activity of the polynucleotide, and Kidins220 is
 useful for identifying an agent which modulates activity of Kidins220,
 protein kinase D (PKD) and the interaction between PKD with Kidins220.
 Kidins220 or its polynucleotide is useful in medicine, e.g. using gene
 therapy or for diagnosing or monitoring progression of cancer in a
 patient. The agent is useful in manufacture of medicament, for use in
 treatment of neurodegenerative disease, and the agent which inhibits
 function of Kidins220 gene or its product is useful for treating cancer
 which is glioblastoma multiforme or prostate cancer. The compound is

CC useful for imaging cancer in an individual, and for diagnosing or
 CC prognosing, and also for treating an individual. The present
 CC sequence represents Rat Kidins220.

SQ Sequence 1763 AA:

Query Match 93.4%; Score 8301; DB 23; Length 1763;
 Best Local Similarity 90.9%; Pred. No. 0;
 Matches 1603; Conservative 47; Mismatches 64; Indels 50; Gaps 2;

QY 1 MSVLISQSVYINVEEBNIPALKALLKCKDKDERNECCQTPMLAEGONEIYKELIKN 60
 DB 1 MSVLISQSVYINVEEBNIPALKALLKCKDKDERNECCQTPMLAEGONEIYKELIKN 60
 QY 61 GANCNLEDDNNFTALISKEGHVHVEELKCGVNLHRDGMGTALMMAKCYKGRTPV 120
 DB 61 GANCNLEDDNNFTALISKEGHVHVEELKCGVNLHRDGMGTALMMAKCYKGRTPV 120
 QY 121 ELLSHGANPSVTGLQYSYVPIIMAAGRHADIHLLQNGAKVNCSDRYGTPPYMAAR 180
 DB 121 ELLSHGANPSVTGLQYSYVPIIMAAGRHADIHLLQNGAKVNCSDRYGTPPYMAAR 180
 QY 181 KGHLECVKHLAMGADVDEGANSMTALIVAKGYTOSVKEILKRNPNVLTJDKGNTA 240
 DB 181 KGHLECVKHLAMGADVDEGANSMTALIVAKGYTOSVKEILKRNPNVLTJDKGNTA 240
 QY 241 LMIAKESGHEITVQDLDAAGTVYNI PDRSGDVLVGAARGHVEIVRALLOKYADIDRG 300
 DB 241 LMIAKESGHEITVQDLDAAGTVYNI PDRSGDVLVGAARGHVEIVRALLOKYADIDRG 300
 QY 301 QDNKTALYAVAEKGNATWTRDILQCNPDTEICTKDETPLIKATKMRNIEVEELLDDKGA 360
 DB 301 QDNKTALYAVAEKGNATWTRDILQCNPDTEICTKDETPLIKATKMRNIEVEELLDDKGA 360
 QY 361 KYSAVDKKCDTPLHAIIRGRSKRLAELIRNPKDGRLLYRPNKAGETPYNIDCSHQKIL 420
 DB 361 KYSAVDKKCDTPLHAIIRGRSKRLAELIRNPKDGRLLYRPNKAGETPYNIDCSHQKIL 420
 QY 421 TQIFGARHLSPTETGDMGVDLYSSALADILSEPTMOPICVGYLAQWGSKSPFLKKL 480
 DB 421 TQIFGARHLSPTETGDMGVDLYSSALADILSEPTMOPICVGYLAQWGSKSPFLKKL 480
 QY 481 EDEMTFFAQGLEPFOFSWMLIVFTLLLCGLGLLFFATVPHNIGIAVSLTALITXF 540
 DB 481 EDEMTFFAQGLEPFOFSWMLIVFTLLLCGLGLLFFATVPHNIGIAVSLTALITXF 540
 QY 541 FTVIYFGGRBEESNMWAVLSTRLARHIGYELLKIMFVNPPLPBOGTKALPVRFLF 600
 DB 541 FTVIYFGGRBEESNMWAVLSTRLARHIGYELLKIMFVNPPLPBOGTKALPVRFLF 600
 QY 601 TTYNRLSSVGGRTSLAEMTATSDACERFGLATRLRFPVFTEDTOGKKKKKRCCLPS 660
 DB 601 TTYNRLSSVGGRTSLAEMTATSDACERFGLATRLRFPVFTEDTOGKKKKKRCCLPS 660
 QY 661 FVIFFLITGCIISITLILAIFFVDPKHLTVNVLISIASVVGAFVLNCRPTMVOYLDL 720
 DB 661 FVIFFLITGCIISITLILAIFFVDPKHLTVNVLISIASVVGAFVLNCRPTMVOYLDL 720
 QY 721 NSQRRRLNASKLHLKLSSEGFMYKCEVELAMARMAKTIDSFTQNGTRLVYIIDGLAC 780
 DB 721 NSQRRRLNASKLHLKLSSEGFMYKCEVELAMARMAKTIDSFTQNGTRLVYIIDGLAC 780
 QY 781 EDDKVLQMDTVYRVLFSKGPFAIRASPHIITIKINONLSVLDNSINIGHDWRNIVH 840
 DB 781 EDDKVLQMDTVYRVLFSKGPFAIRASPHIITIKINONLSVLDNSINIGHDWRNIVH 840
 QY 841 LRVPLNSRGLSNARFLVTSATNGDVPSCDTTQIODEADRRYSONSGEMTKLSGKTALN 900
 DB 841 LRVPLNSRGLSNARFLVTSATNGDVPSCDTTQIODEADRRYSONSGEMTKLSGKTALN 900
 QY 901 RRDYTRRROMORTITRQMSFDLTKLLVIEDWFSIDSPOTMRRLNIVSVTRGLLRANOIS 960
 DB 901 RRDYTRRROMORTITRQMSFDLTKLLVIEDWFSIDSPOTMRRLNIVSVTRGLLRANOIS 960

QY 961 FNMRLASWINTTEOWPYRTSNLLIYEETEGIPDQMLTKTYERISKNIPTTKDVEPL 1020
 DB 961 FNMRLASWINTTEOWPYRTSNLLIYEETEGIPDQMLTKTYERISKNIPTTKDVEPL 1020
 QY 1021 EIDGIRNPEVLSRTPTVLAROVKFLPCTVNLDPRLRELIADVARRAROIISGGLAY 1080
 DB 1021 EIDGIRNPEVLSRTPTVLAROVKFLPCTVNLDPRLRELIADVARRAROIISGGLAY 1080
 QY 1081 PPLPLHEGPPRAPSGYSGPPSVCSSTSPNGFPAGGVSPQPHSSYSGMTGPQHPFYNR- 1139
 DB 1081 PPLPLHEGPPRAPSGYSGPPSVCSSTSPNGFPAGGVSPQPHSSYSGMTGPQHPFYNR- 1139
 QY 1140 -----GSGPAGPVVLL 1151
 DB 1141 FPAPLYTPRYPPGSOHLISRSYKASLPRDQNNGLPCDSGFKRGSRPCHRLQVLT 1200
 QY 1152 NSLNDVAVCEKLEKJEGDQSMLPQYCTTIKKANINGRYLAQCNIDELKTKEMNNFGDMH 1211
 DB 1201 EPNDRGCRCEKLRQJEGDQSMMPQYCTTIKKANINGRYLAQCNIDELKTKEMNNFGDMH 1260
 QY 1212 LFRSTVLEMRNAESHVPEDEPRLFSESSSGAPHPGEPARRASHNELPTELSSQTPYTLN 1271
 DB 1261 LFRSTVLEMRNAESHVPEDEPRLFSESSSGAPHPGEPARRASHNELPTELSSQTPYTLN 1260
 QY 1272 FSEELNLTGLDEGAPRHSNLSMQSOTRRTPSLSLNSQDSIEISKITDKVOAEYRDAY 1331
 DB 1321 FSEELNLTGLDEGAPRHSNLSMQSOTRRTPSLSLNSQDSIEISKITDKVOAEYRDAY 1380
 QY 1332 REYIAQMSOLEGGPSTTISGRSPHSTYMGQSSGSIHNSLEQEKGDSEKPPDDGR 1391
 DB 1381 REYIAQMSOLEGGPSTTISGRSPHSTYMGQSSGSIHNSLEQEKGDSEKPPDDGR 1391
 QY 1392 KSFLLKRGDYIDYSSSVSTNDASPLDYTEDEKSDQSGSKLLPGKSSERSSTFQDNL 1451
 DB 1441 KSFLLKRGDYIDYSSSVSTNDASPLDYTEDEKSDQSGSKLLPGKSSERSSTFQDNL 1500
 QY 1452 KFKGSLRGYOKLPSEDESGTESDNTPLKDDKRAKAEYVERPKSPREHSAEPIRTFI 1511
 DB 1501 KFKGSLRGYOKLPSEDESGTESDNTPLKDDKRAKAEYVERPKSPREHSAEPIRTFI 1560
 QY 1512 KAKYELSDALDKDSDSGVRSSESSPNHSLHNEVADDSQLEKANLIELEDDSHSGRG 1571
 DB 1561 KAKYELSDALDKDSDSGVRSSESSPNHSLHNEVADDSQLEKANLIELEDDSHSGRG 1620
 QY 1572 IPHSLSGIDPITIAMSTICSEDKKSPSCSLIASSPEENWPAQCAAYMLNTPSTVTLNN 1631
 DB 1621 IPHSLSGIDPITIAMSTICSEDKKSPSCSLIASSPEENWPAQCAAYMLNTPSTVTLNN 1680
 QY 1632 NSAPANRANONDEMEGIRETSQVILRPSSESPNPTTINENILKMTHTRSORSSTYRLSK 1691
 DB 1681 NSAPANRANONDEMEGIRETSQVILRPSSESPNPTTINENILKMTHTRSORSSTYRLSK 1740
 QY 1692 DPPELHAAASSESTGFGERESITL 1715
 DB 1741 DASELH-AAASSESTGFGERESITL 1763

RESULT 6

ID AAG75604 standard; Protein; 705 AA.

AAG75604;

03-SEP-2001 (first entry)

Human colon cancer antigen protein SEQ ID NO:6368.

Human colon cancer; colon cancer antigen; diagnosis; detection;

colorectal carcinoma.

Homo sapiens.

XX

PN W0200122920-A2.
 XX 05-APR-2001.
 PD 28-SEP-2000; 2000MO-US26524.
 XX 29-SEP-1999; 9905-0157137.
 PR 03-NOV-1999; 9905-0163280.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA Ruben SM, Barash SC, Birse CE, Rosen CA;
 PI WPI, 2001-235357/24.
 DR N-PSDB; AAH35009.
 XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
 PT useful for preventing, diagnosing and/or treating colorectal cancers -
 PS Claim 11; Page 7830-7834; 9803pp; English.
 XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
 CC cancer-associated nucleic acid molecules (N) and proteins (P), where
 CC the proteins are collectively known as colon cancer antigens. The colon
 CC cancer antigens have cytostatic activity and can be used in gene
 CC therapy and vaccine production. N and P may be used in the prevention,
 CC diagnosis and treatment of diseases associated with inappropriate P
 CC expression. For example, N and P may be used to treat disorders
 CC associated with decreased expression by rectifying mutations or deletions
 CC in a patient's genome that affect the activity of P by expressing P.
 CC Inactive proteins or to supplement the patients own production of P.
 CC Additionally, N may be used to produce the colon cancer-associated Ps,
 CC by inserting the nucleic acids into a host cell and culturing the cell
 CC to express the proteins. N and P can be used in the prevention, diagnosis
 CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
 CC and AAB77789 represent sequences used in the exemplification of the
 CC present invention.
 CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
 CC missing at time of publication, meaning no sequences are present for
 CC SEQ ID NO:1027 to 1052, 7921 and 7922.
 XX
 SQ Sequence 705 AA;
 Query Match 34.3%; Score 3050.5; DB 22; Length 705;
 Best Local Similarity 98.8%; Pred. No. 1e-234; 4; Indels 1; Gaps 1;
 Matches 587; Conservative 2; Mismatches 4;
 QY 62 ANCNLELDLNMWTLISASKEGHVHVEELKCGVNEHRDMGWTALMACYKGRDVE 121
 DB 28 SNCNLELDLNMWTLISASKEGHVHVEELKCGVNEHRDMGWTALMACYKGRDVE 87
 QY 122 LLLSHGNPSVNTGLQSVYPTIWAAGRGADIVHLLONGAKYNCSDKGTPIVAAAR 181
 DB 88 LLLSHGNPSVNTGLQSVYPTIWAAGRGADIVHLLONGAKYNCSDKGTPIVAAAR 146
 QY 182 GHLECYKHLAMGADVDQEGANSMTALIVAVKGYTQSYKEILKRPNVNLTDKGNATL 241
 DB 147 GHLECYKHLAMGADVDQEGANSMTALIVAVKGYTQSYKEILKRPNVNLTDKGNATL 206
 QY 242 MIASKEGHEIYODLLDAGTYVNIIPRSGDTVLIGAVGHEIVALLQKKADIDIRQ 301
 DB 207 MIASKEGHEIYODLLDAGTYVNIIPRSGDTVLIGAVGHEIVALLQKKADIDIRQ 266
 QY 302 DNKTALYMAEKNAMVNRDILQCNPDTECTKDGTEPLIKATKMNIEVVELLDKGA 361
 DB 267 DNKTALYMAEKNAMVNRDILQCNPDTECTKDGTEPLIKATKMNIEVVELLDKGA 326
 QY 362 VSAVADKKGDPPLAIALRGRSRKLAELLRNPKDGRLLYPNKAETPYNIDCSHOKSILT 421
 DB 327 VSAVADKKGDPPLAIALRGRSRKLAELLRNPKDGRLLYPNKAETPYNIDCSHOKSILT 386
 QY 422 QIFGARHLSPTEEDGMDLGYDLYSSALADILSEPTMOPICVGLVAMQSSGKSFLLKLE 481
 DB 481 QIFGARHLSPTEEDGMDLGYDLYSSALADILSEPTMOPICVGLVAMQSSGKSFLLKLE 441

Db 387 QIGFARHLSPTETDGMLEGYDYSALADILSEPTMQPVCVLYAQMGSGKSLKLE 446
 QY 482 DEKTRAGQOIEPLFOFSLIVFTLLCGGLFAFVHPMLGAVLSFLALLYTF 541
 Db 447 DEKTRAGQOIEPLFOFSLIVFTLLCGGLFAFVHPMLGAVLSFLALLYTF 506
 QY 542 IVIFGGRGSGESMNNAMVLSRLARHIGYLELLKLMVNPPELPEQTKALPVAFLE 601
 Db 507 IVIFGGRGSGESMNNAMVLSRLARHIGYLELLKLMVNPPELPEQTKALPVAFLE 566
 QY 602 DYNRLSSVGETSLAEMIALTSDACEREGFLATRLPVFKTEDTGKKKKKT 655
 Db 567 DYNRLSSVGETSLAEMIALTSDACEREGFLATRLPVFKTEDTGKKKKKT 620

RESULT 7
 AAE01035 standard; Protein: 551 AA.
 AAE01035:
 04-JUL-2001 (first entry)

Human death domain-containing receptor (DDCR) protein from HODDX59 clone.
 Human: death domain-containing receptor; DDCR; immunosuppressive;
 antiathritic; antirheumatic; antiproliferative; cyostatic;
 cardiant; vasotropic; cerebroprotective; neurotropic; neuroprotective;
 antimicrobial; virucide; fungicide; ophthalmological; gene therapy;
 immunodeficiency disease; Acquired immune deficiency syndrome;
 AIDS; leukaemia; autoimmune disease; systemic lupus erythematosus;
 hyperproliferative disorder; neoplasm; cerebrovascular disorder;
 cerebral ischaemia; angiogenesis; cardiovascular disorder;
 neurodegenerative disease; Alzheimer's disease; Parkinson's disease;
 ocular disorder; corneal infection; degenerative disease; SMA; apoptosis;
 spinal muscular atrophy; epithelial cell proliferation; infection;
 cancer; wound healing; skin aging; chemotaxis; HODDX59 clone.

OS Homo sapiens.
 XX
 XX
 FH Key Location/Qualifiers
 FT Domain 399..415
 FT /label=Transmembrane_domain
 FT 426..442
 FT /label=Transmembrane_domain
 FT 64..69
 FT /label=Immunogenic_epitope
 FT 131..136
 FT /label=Immunogenic_epitope
 FT 164..169
 FT /label=Immunogenic_epitope
 FT 197..203
 FT /label=Immunogenic_epitope
 FT 229..235
 FT /label=Immunogenic_epitope
 FT 298..316
 FT /label=Immunogenic_epitope
 FT 328..335
 FT /label=Immunogenic_epitope
 FT 445..452
 FT /label=Immunogenic_epitope
 FT 482..490
 FT /label=Immunogenic_epitope
 FT 540..551
 FT /label=Immunogenic_epitope
 FT Region
 XX WO200129063-A2.
 XX 26-APR-2001.
 XX 17-OCT-2000; 2000WO-US28666.
 XX 18-OCT-1999; 9905-0159585.

PR 24-NOV-1999; 9905-0167246.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Ni J, Ruben SM;
 XX
 DR WPI: 2001-300316/31.
 XX N-PSDB: AAD04776.
 PT New death domain containing receptor polynucleotides and polypeptides,
 PS useful for treating and diagnosing cancer
 XX
 PS Claim 1; Page 292-294; 298pp; English.

The patent discloses novel death domain-containing receptor (DDCR)
 CC cDNAs and their corresponding proteins. DDCR cDNA and protein are
 CC used to prevent, treat or ameliorate a medical condition in mammals.
 CC They are also used in diagnosing a pathological condition or
 CC susceptibility to a pathological condition. The DDCR protein and
 CC its antibodies are used in the diagnosis and treatment of disorders
 CC such as immunodeficiency diseases (e.g. Acquired immune deficiency
 CC syndrome (AIDS), leukaemia) autoimmune diseases (e.g. systemic lupus
 CC erythematosus, rheumatoid arthritis), hyperproliferative disorders
 CC (e.g. neoplasms of the breast or liver), cerebrovascular disorders
 CC (e.g. cerebral ischaemia, angiogenesis), cardiovascular disorders
 CC (e.g. cardiac arrest), neurodegenerative diseases (e.g. Alzheimer's
 CC disease, Parkinson's disease), ocular disorders (e.g. corneal
 CC infection), degenerative diseases (e.g. spinal muscular atrophy-SMA),
 CC cancer, aberrant apoptosis, disorders of the placenta or uterus and
 CC infections caused by bacteria, viruses and fungi. The DDCR proteins
 CC are used to aid wound healing and epithelial cell proliferation,
 CC to prevent skin aging due to sunburn, to maintain organs before
 CC transplantation, to support cell culture of primary tissues, to
 CC regenerate tissues and in chemotaxis. The DDCR sequences are used
 CC in correcting aberrant cellular apoptosis by gene therapy.
 CC The present sequence is human death domain-containing receptor
 CC (DDCR) protein from clone HODDX59.
 XX
 XX Sequence 551 AA:

Query Match 32.1%; Score 2851.5; DB 22; Length 551;
 Best Local Similarity 99.6%; Pred. No. 5.7e-219;
 Matches 550; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 102 MGSMTALMAMACYGRTDVVELLSHGANSYTGLOYSYPIIMAGRGHADVHLLONG 161
 Db 1 MGSMTALMAMACYGRTDVVELLSHGANSYTGLOYSYPIIMAGRGHADVHLLONG 59
 QY 162 AKVNSDKYGTTPVLAARKGHILECYKHLAMGADYDVGANSMTALYAVAGGYTSVK 221
 Db 60 AKVNSDKYGTTPVLAARKGHILECYKHLAMGADYDVGANSMTALYAVAGGYTSVK 119
 QY 222 EILKRNPNVNLNDKDNLTALMTASKEGHEIYODLLDAGTYNIPRSGDYVLIGAVRG 281
 Db 120 EILKRNPNVNLNDKDNLTALMTASKEGHEIYODLLDAGTYNIPRSGDYVLIGAVRG 179
 QY 282 HVEIVALLQKYADIDIRGDNKTALYMAVEKGNATMVBDILQCNDETEICRDETPLI 341
 Db 180 HVEIVALLQKYADIDIRGDNKTALYMAVEKGNATMVBDILQCNDETEICRDETPLI 239
 QY 342 KATKRNIEVELLDKGAKSAVDKGGTPLHAIKRSKRIALLNPNDRGLYRP 401
 Db 240 KATKRNIEVELLDKGAKSAVDKGGTPLHAIKRSKRIALLNPNDRGLYRP 299
 QY 402 NKAGETPNYNDGSHOKSIITQIFGARHLSPTETDGMLEGYDYSALADILSEPTMQPPI 461
 Db 300 NKAGETPNYNDGSHOKSIITQIFGARHLSPTETDGMLEGYDYSALADILSEPTMQPPI 359
 QY 462 CVGLYQMGSGKSFLLKLEDEKTRAGQOIEPLFOFSLIVFTLLCGGLFAFV 521
 Db 360 CVGLYQMGSGKSFLLKLEDEKTRAGQOIEPLFOFSLIVFTLLCGGLFAFV 419
 QY 522 HPNLGIAVLSFLALLYFIYIYIFGGRGSGESMNNAMVLSRLARHIGYLELLKLMV 581

```

DB 420 HNNLGLAVSLFLALLIYFIYITGGRREGSMMWAWLSTRLARHIGYELLKLMFV 479
OY 582 NPPELPEQOTKALPVRFLETDYNNRLSVGGESTLAEMTATLSDACEREGFATRLFRVF 641
DB 480 NPPELPEQOTKALPVRFLETDYNNRLSVGGESTLAEMTATLSDACEREGFATRLFRVF 539
OY 642 KTEPDQGGKKKK 653
DB 540 KTEPDQGGKKKK 551

RESULT 8
AAB95191
ID AAB95191 standard; Protein; 543 AA.
AC AAB95191;
DT 26-JUN-2001 (first entry)
XX
DE Human protein sequence SEQ ID NO:17273.
OS Homo sapiens.
XX
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
XX Homo sapiens.
XX
XX EP1074617-A2.
XX
XX 07-FEB-2001.
XX
XX 28-JUL-2000; 2000EP-0116126.
XX
XX 29-JUL-1999; 99JP-0248036.
XX
XX 27-AUG-1999; 99JP-0300253.
XX
XX 11-JAN-2000; 2000JP-0118776.
XX
XX 02-MAY-2000; 2000JP-0183767.
XX
XX 09-JUN-2000; 2000JP-0241899.
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
XX full-length cDNAs defined in the specification, and for the detection
XX and/or diagnosis of the abnormality of the proteins encoded by the
XX full-length cDNAs -
XX
XX Claim 8; SEQ ID 17273; 2537bp + CD ROM; English.
XX
XX The present invention describes primer sets for synthesizing 5602
XX full-length cDNAs defined in the specification. Where a primer set
XX comprises: (a) an oligo-dr primer and an oligonucleotide complementary
XX to the complementary strand of a polynucleotide which comprises one of
XX the 5602 nucleotide sequences defined in the specification, where the
XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX of an oligonucleotide comprising a sequence complementary to the
XX complementary strand of a polynucleotide which comprises a 5'-end
XX sequence and an oligonucleotide comprising a sequence complementary to a
XX polynucleotide which comprises a 3'-end sequence, where the
XX oligonucleotide comprises at least 15 nucleotides and the combination of
XX the 5'-end sequence/3'-end sequence is selected from those defined in
XX the specification. The primer sets can be used in antisense therapy and
XX in gene therapy. The primers are useful for synthesizing polynucleotides,
XX particularly full-length cDNAs. The primers are also useful for the
XX detection and/or diagnosis of the abnormality of the proteins encoded by
XX the full-length cDNAs. The primers allow obtaining of the full-length
XX cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
XX AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
XX AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
XX represent oligonucleotides, all of which are used in the exemplification

```

```

CC of the present invention.
XX
XX Sequence 543 AA;
OY Query Match 31.4%; Score 2791; DB 22; Length 543;
DB Best Local Similarity 99.6%; Pred. No. 3.9e-214;
OY Matches 541; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY 1173 MLPQCTTIKKANINGVLAQCNIIDELKRMNNFGDMHFRSTVLEMRNASHVYVPEP 1232
DB 1 MLPQCTTIKKANINGVLAQCNIIDELKRMNNFGDMHFRSTVLEMRNASHVYVPEP 60
OY 1233 RFLSESSSGPAPGEPARRASHNELPHTELSSQTPYTLNFSFEELNTLGLDEGAPRHSNL 1292
DB 61 RFLSESSSGPAPGEPARRASHNELPHTELSSQTPYTLNFSFEELNTLGLDEGAPRHSNL 120
OY 1293 SWSQTRRTPSLSSLSNSQDSISITKTDVQVQAEYDAVREYIAQMSOLEGGGSTTISG 1352
DB 121 SWSQTRRTPSLSSLSNSQDSISITKTDVQVQAEYDAVREYIAQMSOLEGGGSTTISG 180
OY 1353 RSSPHSTYYMGQSSGSIHSNLEQKRGKDEPKPDGGRKSPFLMRGVDYIDYSSGCVSTN 1412
DB 181 RSSPHSTYYMGQSSGSIHSNLEQKRGKDEPKPDGGRKSPFLMRGVDYIDYSSGCVSTN 240
OY 1413 DASPLDPTFEDEKSDQSGSKLLPGKSSERSSLFQTDILKSGSLRYQKLPSEDESGT 1472
DB 241 DASPLDPTFEDEKSDQSGSKLLPGKSSERSSLFQTDILKSGSLRYQKLPSEDESGT 300
OY 1473 EESDNTPLKDDDRKAEKGVVERVPKSPESAPRTFTIKAREYLSDALDKKSDSGV 1532
DB 301 EESDNTPLKDDDRKAEKGVVERVPKSPESAPRTFTIKAREYLSDALDKKSDSGV 360
OY 1533 RSSESSPNHSLHNEVADDSOLEKANLIELEDSSHSGRGIPHSLSGLDPIIARMSICSE 1592
DB 361 RSSESSPNHSLHNEVADDSOLEKANLIELEDSSHSGRGIPHSLSGLDPIIARMSICSE 420
OY 1593 DKSPSECSLIASPEENPACOKAVNLRFTSTVTLNNSAPANRANONFDEMGIRET 1652
DB 421 DKSPSECSLIASPEENPACOKAVNLRFTSTVTLNNSAPANRANONFDEMGIRET 480
OY 1653 SOVILPSSPNPTTTONENIKSMTHKRSORSSTYPLSDPELHAASSESTGGEERE 1712
DB 481 SOVILPSSPNPTTTONENIKSMTHKRSORSSTYPLSDPELHAASSESTGGEERE 540
OY 1713 SIL 1715
DB 541 SIL 543

RESULT 9
AAE03645
ID AAE03645 standard; Protein; 513 AA.
XX
XX AAE03645;
DT 06-AUG-2001 (first entry)
XX
XX Human extracellular matrix and cell adhesion molecule-9 (XMA9-9).
XX
XX Human; extracellular matrix and cell adhesion molecule: XMA9.
XX gene therapy; genetic disorder; adrenoleukodystrophy; leukaemia;
XX Down's syndrome; cystic fibrosis; Gaucher's disease; myotonic dystrophy;
XX sickle cell anaemia; thalassaemia; autoimmune disorder; adenocarcinoma;
XX inflammatory disorder; acquired immune deficiency syndrome; AIDS;
XX Addison's disease; allergy; anaemia; asthma; atherosclerosis; melanoma;
XX Crohn's disease; diabetes mellitus; atopic dermatitis; lymphoma; cancer;
XX glomerulonephritis; multiple sclerosis; Grave's disease; osteoarthritis;
XX osteoporosis; psoriasis; rheumatoid arthritis; ulcerative colitis;
XX infection; cell proliferative disorder; actinic keratosis; myeloma;
XX arteriosclerosis; neotrophic; anticonvulsant; antithyroid; nephrotropic;
XX neuroprotective; dermatological.
XX
XX Homo sapiens.
XX
XX OS

```

XX Key Location/Qualifiers
 FH Region 94..108
 FT Domain /note="Adrenocorticotrophin receptor"
 FT 196..198
 FT /note="Rgd cell interaction motif"
 FT Region 328..335
 FT /note="Phage integrase family"
 XX W0200142285-A2.
 XX 14-JUN-2001.
 PD 05-DEC-2000; 2000WO-US32990.
 PF 10-DEC-1999; 99US-0172852.
 PR 16-DEC-1999; 99US-0172354.
 XX (INCY-) INCYTE GENOMICS INC.
 XX Yue H, Tang YT, Lal P, Burford N, Azimzai Y, Patterson C,
 PI Baughn MR, Lu DM, Shah P, Au-Young J;
 DR WPI: 2001-381632/40.
 DR N-PSDB: AAD08053.
 XX New human extracellular matrix and cell adhesion molecules and
 PT polynucleotide sequences encoding them, useful for diagnosis,
 PT prevention, treatment of genetic, autoimmune and cell proliferative
 PT disorders
 PS Claim 1; Page 101-102; 135pp; English.
 XX The present sequence is a human extracellular matrix and cell
 CC adhesion molecule (XMAP). The XMAP is used for screening a compound for
 CC effectiveness as an agonist or antagonist of XMAP. The identified agonist
 CC or antagonist are used for treating a disease or condition associated
 CC with decreased or increased expression of functional XMAP. The
 CC polynucleotides encoding XMAP are useful in somatic or germline gene
 CC therapy to correct a genetic deficiency, to express a conditionally
 CC lethal gene product and to express a protein which affords protection
 CC against intracellular parasites and also for diagnosis of disorders
 CC associated with expression of XMAP. They are also used for generating
 CC hybridisation probes useful in mapping the naturally occurring genomic
 CC sequences and to create knock in humanised animals (pigs) or transgenic
 CC animals (mice or rats) to model human diseases. Oligonucleotide or longer
 CC fragments derived from the polynucleotide sequences may be used as
 CC elements on a microarray. Antibodies which specifically bind XMAP may be
 CC used for the diagnosis of disorders associated with the expression of
 CC XMAP, or in assays to monitor patients being treated with XMAP. Diseases
 CC diagnosed, prevented or treated include genetic disorders such as
 CC adrenoleukodystrophy, Down's syndrome, cystic fibrosis, Gaucher's
 CC disease, myotonic dystrophy, sickle cell anaemia, thalassemia,
 CC autoimmune/inflammatory disorders such as acquired immune deficiency
 CC syndrome (AIDS), Addison's disease, allergies, anaemia, asthma,
 CC atherosclerosis, Crohn's disease, diabetes mellitus, atopic dermatitis,
 CC glomerulonephritis, multiple sclerosis, Grave's disease, osteoarthritis,
 CC osteoporosis, psoriasis, rheumatoid arthritis, ulcerative colitis,
 CC bacterial, fungal, parasitic, protozoal and helminthic infections and
 CC cell proliferative disorders such as actinic keratosis, arteriosclerosis
 CC and cancer including breast, bladder, bone marrow, brain and uterus
 CC cancer, leukaemia, adenocarcinoma, lymphoma, melanoma and myeloma.

Sequence 513 AA;

Query Match 29.7%; Score 2638; DB 22; Length 513;
 Best Local Similarity 99.8%; Pred. No. 6.3e-202;
 Matches 512; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1203 MNMFGDWHLPFRSTVLEMRNESHVPEPDPFLSESSSGPAPHPGEPARRASHNLPTEL 1262
 DB 1 MNMFGDWHLPFRSTVLEMRNESHVPEPDPFLSESSSGPAPHPGEPARRASHNLPTEL 60

QY 1263 SSQTPYTLNFSPEELNTGLDEGAPRHSNLSWQSOTRRTPSLSLNSQSSLEISKLDDK 1322
 DB 61 SSQTPYTLNFSPEELNTGLDEGAPRHSNLSWQSOTRRTPSLSLNSQSSLEISKLDDK 120
 QY 1323 VOAEYRDAYREYIAQMSOLEGGPGSTTISGRSSPHSTYMGQSSGSGSIHNSLEDEKGD 1382
 DB 121 VOAEYRDAYREYIAQMSOLEGGPGSTTISGRSSPHSTYMGQSSGSGSIHNSLEDEKGD 180
 QY 1383 SEPKPDDGKRSFLMRGVDYIDSSGVSFNDA SPLDPTDEDEKSDQSGSKLPEKKSSE 1442
 DB 181 SEPKPDDGKRSFLMRGVDYIDSSGVSFNDA SPLDPTDEDEKSDQSGSKLPEKKSSE 240
 QY 1443 RSLFOTDLKLGSGLRQKLPSPDEDESGTEESDNTPLIKDDKDKAKGVEYRPPKSPDH 1502
 DB 241 RSLFOTDLKLGSGLRQKLPSPDEDESGTEESDNTPLIKDDKDKAKGVEYRPPKSPDH 300
 QY 1503 SAEPITFTKAEYVSDALIDKDDSDSGVRSSESPNHSILNEYADSOLEKANLILE 1562
 DB 301 SAEPITFTKAEYVSDALIDKDDSDSGVRSSESPNHSILNEYADSOLEKANLILE 360
 QY 1563 DSHSGKRGIPHSLSGLDPIIARMSICSEDKKSPSECSLIASPEENWPACOKAYNLNR 1622
 DB 361 DSHSGKRGIPHSLSGLDPIIARMSICSEDKKSPSECSLIASPEENWPACOKAYNLNR 420
 QY 1623 TPSTVTLNNSNAPANRANONFDEMGIRETSQVILRPSSSPPTTIONENLKSMTHKRSQ 1682
 DB 421 TPSTVTLNNSNAPANRANONFDEMGIRETSQVILRPSSSPPTTIONENLKSMTHKRSQ 480
 QY 1683 RSYTYRLSKDPPELHAAASSESTGGEKEESTL 1715
 DB 481 RSYTYRLSKDPPELHAAASSESTGGEKEESTL 513

RESULT 10

ABG08697
 ID ABG08697 standard; Protein; 624 AA.

XX ABG08697;

AC 13-FEB-2002 (first entry)

XX Novel human diagnostic protein #8688.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

KM food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX W0200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001MO-US08631.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Dmanac RT, Liu C, Tang YT.

XX WPI: 2001-639362/73.

XX N-PSDB: AAS72884.

XX New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PS biodegradability
 Claim 20; SEQ ID NO 39056; 103pp; English.
 The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,


```

QY 420 LTOIGARHLSPETDGMGLDYLYSSALADILSEPTMOPICGVGYAOMSGSKSFLKK 479
D 442 LGVYGARLINTNDESGMLYELYSALADVLSEPTLTPIVGLYAKMGSKSFLKK 501
QY 480 LEDEMKTFAGQOIEPLFQFSMLVELT---LLCG---GIGLFAFVHNLGIANVLSF 533
D 502 LREMNNAFARQMAEPPIRTSGLLFVCLHVALLIGIVGISTWSAV---VGVSAVGF 556
QY 534 LALLYFIVFYCGREGESGEMWAWLSTRARHIGYELLLKMFVNPPELPEQOTRK 593
D 557 LLLAYLLLAARVCNVO---MDMOMAVSVQGLEKRMRLILIQVAFCHPG---PQSDQA 613
QY 594 LPVAFLEFDYNRSLSSVGGETSLAEMIAITLSDACEREFGLATRLFRVFKTE---DTQKKR 651
D 614 KPVAFHFAEANSASPT---GDGAVAHMLAALLDAIESHVGMATRLRYAFRCKLKVDMGR 672
QY 652 MKKTCCLPSFVIFLFIIGIISGTTLAIF---RDPKHLVYNAVLIASVGLAFV 706
D 673 WRRKCCIPYVLELALVAVTGTISLVAFTFADEKEKHILV---ALVYIAAVMGTLIC 730
QY 707 LNCRTMVOVLDSLINSQRKRLHNAASKLHKLKSGEFKVLKCEVELMARAKTIDSPTON 766
D 731 TLHLVLAKEVVSLETSIRVLAKEAV---RSSASAPLTMGAIVAVMTMVKCLDAFTNQ 786
QY 767 QTRLVVIIDGLDACEODKVLQMDTVKVLFS---KGPEIALPASPDLIIIRAINQNLNVL 824
D 787 QSRIVGVIDALDSCDERILTLINAVQTLSSPFRPVLIIISDPHVIKAAABANSRRL 846
QY 825 RDSINNGHDMRNIVHLPVFLNRSGLSNARKFLVTS---AINGVPCSD---TTGQED 877
D 847 TEGIGGHDFLRNLVHLPVYLQNSGLKRVQRAQMTALLFRSGGDDQTDGPTLIGSVS 906
QY 878 ADDRVSQNS---LGEMLTL-----GSKTALN-----RQTYRRMQRTITQMSF 920
D 907 A-RRLSNASELISQEKLRPARGGGKRLRLSESVASGSLNHLRGQPVY-----L 960
QY 921 DLKRLVTEPMFSDISQYMRRLNIVSVTGLRLANQISFNDRLASWINTLEQWYRT 980
D 961 DLSRIVLTIDYFSDVNPERSMRRLMNVYIVTRLLKAFQIEFWSYRLSSWINTLEQWPLRA 1020
QY 981 SWILVLEE---TEGIPQMTLKIYERISKNIPTTKVEBLELDGDIRFVEFLSRT 1038
D 1021 SMYLHNDQFMDSNADSVSLQSYTEKLRKLAVALLEDRERKLDAFLQHK 1080
QY 1039 VLVAADVKEFLPCTVNLDPKRLRETIAD-----VRAREQISIGGLAAPPPLHE 1087
D 1081 DLVAADRLRFLPFTINLDPLRKVLAKEDQFTIEDEGLVIAQARPSVNTMRQP----- 1134
QY 1088 GPPAPSGYQSPESVCSSTFNGPFGAGVYSPQPHSSY-----YSGMTG 1131
D 1135 ---APPTY-----VSPQAYPYQMFQONEYPANELRSNLTSTTE 1171
QY 1132 PQHPFYNRGSGPAPGPVY---LNLNLANVAVCEKIKQIEGDQSLPQYCTIKKANNGR 1189
D 1172 PVTPPLINSPDSREDDILQTKLDLYEGYISLDRJEDM-KRALPKLAVYLENAINGR 1230
QY 1190 VLAOCNIDELKKEKNNMNGDMHLEFRSTVLEMRNAESHVVP-----EDPRFSESSGP 1242
D 1231 VLKHCMDPLKSVGLSFGHMELEFRLLITTLREOER-LPRKQOQOQCGALAEAPSNVP 1288
QY 1243 APHG-----EPAR---ASHNELPHE---LSSQ---TPYLLNLSFEELNLTGLDEGA 1286
D 1289 MIKQVTDALMQPRESLSRKNNSVSHMEKQVTEQMCTGLQTLNEAAY-DVAASSERPS 1347
QY 1287 PRHNLNLSMOSOTRRTPLSL-----NSODSSIEISKLTDVQ---AAV-EDAVREYI 1335
D 1348 PTGMLAVALQOLAPLRESSESGSPDDQKQYGVKLSNNNNNOYTHAAYNSVSSHSL 1407
QY 1336 AOMSOLEBGGPGSTTISGRSSPHSTYYMGQSSGSIHNSLNEOEKGD 1382
D 1408 QSLSTLVGAP-----VGHGSGSG---SHLHIGND 1435

```

```

RESULT 12
AAB94785
ID AAB94785 standard; Protein; 129 AA.
XX
AC AAB94785;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human protein sequence SEQ ID NO:15890.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000BP-0116126.
XX
PR 29-JUL-1999; 993P-0248036.
PR 27-AUG-1999; 993P-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELT-) HELIX RES INSE.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
PS Claim 8; SEQ ID 15890; 2537bp + CD ROW; English.
XX
CC The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises at least 15 nucleotides and the combination
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 129 AA;
XX
Query Match 7.5%; Score 666; DB 22; Length 129;
Best Local Similarity 100.0%; Pred. No. 4.5e-45;
Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1587 MSICSEDKKSPSCSLIASSPEENMPACOKAYNLNRFTSTVYLNNSAPANRANONPEM 1646
D 1 MSICSEDKKSPSCSLIASSPEENMPACOKAYNLNRFTSTVYLNNSAPANRANONPEM 60

```

QY 1647 EGIRETSQVILRPSSSPNPTTTONENLKMTHKRSORSSYTRLSKDPPELHAAASSESTG 1706
 |||||||
 Db 61 EGIRETSQVILRPSSSPNPTTTONENLKMTHKRSORSSYTRLSKDPPELHAAASSESTG 120
 QY 1707 FGEERESII 1715
 |||||||
 Db 121 FGEERESII 129

RESULT 13
 AAM40811
 ID AAM40811 standard; Protein: 120 AA.
 AC AAM40811;
 XX 22-OCT-2001 (first entry)
 DT Human polypeptide SEQ ID NO 5742.
 DE Human; nocitropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KM peripheral nervous system; neuropathy; central nervous system; CNS;
 KM Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KM amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KM chemokine; thrombolytic; drug screening; arthritis; inflammation;
 KM leukaemia.
 XX Homo sapiens.
 OS WO20015312-A1.
 XX 26-JUL-2001.
 PD 26-DEC-2000; 2000WO-US34263.
 PF 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.
 XX (HYSE-) HYSEQ INC.
 PA Tang YF, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX WPI: 2001-442253/47.
 XX N-PSDB: AAI59967.
 PT Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 XX Example 2; SEQ ID NO 5742; 10078pp; English.
 PS The invention relates to human nucleic acids (AAI57798-AAI61369) and
 CC the encoded polypeptides (AAM38642-AAM42213) with nocitropic.
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localized neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.

XX SO Sequence 120 AA;
 Query Match 7.0%; Score 620; DB 22; Length 120;
 Best local similarity 100.0%; Pred. No. 1.9e-41;
 Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 217 TOSVKEITLKNPNVNLTKDKNTALMASKEGTEIVODLLDAGTVNIPDSGDTVLIG 276
 |||||||
 Db 1 TOSVKEITLKNPNVNLTKDKNTALMASKEGTEIVODLLDAGTVNIPDSGDTVLIG 60

QY 277 AVRGHVEIYRALLQKADIDIDIGODNKATLYAVKGNATVYRDLQCNPTETCTKDG 336
 |||||||
 Db 61 AVRGHVEIYRALLQKADIDIDIGODNKATLYAVKGNATVYRDLQCNPTETCTKDG 120

RESULT 14
 AAB56277
 ID AAB56277 standard; Protein: 109 AA.
 AC AAB56277;
 XX 13-MAR-2001 (first entry)
 DT Human secreted protein sequence encoded by gene 82 SEQ ID NO:371.
 DE Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;
 KM antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;
 KM cerebroprotective; nocitropic; neuroprotective; antibacterial; vitucide;
 KM fungicide; ophthalmological; gene therapy; pathological condition;
 KM autoimmune disease; rheumatoid arthritis; hyperproliferative disorder;
 KM neoplasm; cardiovascular disorder; cardiac arrest; cerebral ischaemia;
 KM cerebrovascular disorder; angiogenesis; nervous system disorder;
 KM Alzheimer's disease; infection; ocular disorder; corneal infection;
 KM wound healing; skin aging; food additive; preservative.
 XX Homo sapiens.
 OS WO200070042-A1.
 XX 23-NOV-2000.
 PD 11-MAY-2000; 2000WO-US12788.
 PF 13-MAY-1999; 99US-0134068.
 PR (HUMAN-) HUMAN GENOME SCI INC.
 PA Rosen CA, Ruben SM, Moore PA, Young PE, Komatsoulis GA, Birse CE;
 PI Duan RD, Florence KA, Soppet DR;
 XX WPI: 2000-679828/66.
 PT Isolated nucleic acid molecule encoding a human secreted protein is
 PT used in preventing, treating or ameliorating a medical condition -
 XX Disclosure; Page 1025-1026; 1065pp; English.
 PS The polynucleotide sequences given in AAC99818 to AAC99977 encode the
 CC human secreted proteins given in AAB56077 to AAB56362. Human secreted
 CC proteins have activities based on the tissues and cells the genes are
 CC expressed in. Examples of activities include: immunosuppressive;
 CC antirheumatic; antiproliferative; cytostatic; cardiant;
 CC vasotropic; cerebroprotective; nocitropic; neuroprotective; antibacterial;
 CC vitucide; fungicide; and ophthalmological. The human secreted
 CC polynucleotides and proteins can be used to prevent, treat or ameliorate
 CC a medical condition in e.g. humans, mice, rabbits, goats, horses, cats,
 CC dogs, chickens or sheep. They are also used in diagnosing a pathological
 CC condition or susceptibility to a pathological condition. Disorders which
 CC are diagnosed or treated include autoimmune diseases e.g. rheumatoid
 CC arthritis, hyperproliferative disorders e.g. neoplasms of the breast or
 CC liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular
 CC disorders e.g. cerebral ischaemia, angiogenesis, nervous system disorders

Db 824 TTTTTEKHLNVPETMEFVLDSDEGD-DTMTGDGGEYLPEDLKELDGSSLPSSQFLD 882
 QY 635 TRLRVEFTEDTQGGKKKKKCTCCLPYVIFLFIIGCIIISGTLTALTRVDPKHLTVANVL 694
 Db 883 GMYLIRYSLEGGSRDS-----LRSSDSHLSHASY 915
 QY 695 ISIASVGLAVVLCRTWMOY---LDSLNSOR---KRLHNA---SKLHK----- 736
 Db 916 LRDSVMDSVYIISH---QYSTLAKEAKRNSYLSMGTENDLVALSSSIHSGFLVSF 972
 QY 737 -----LKSEGFVKVL--KC--EVELMARMAKTIDFTONOTRLVY---IIDGL 777
 Db 973 MYDAGGAMRGCRHNGRLIIPRKCPTAPTRVTCRLVK-----RHRLATMPMEGE 1024
 QY 778 DACEODKVLQMLDYRVLFSGPFLAIFASDPH---IILKAINQNLNSLRNSIN--GHD 833
 Db 1025 GLASR-----LIEVGPSSAQFLGKHLPTAPPLNEGESLVSRLQLGPPQTK 1072
 QY 834 YMRN-IYHLPEVLSNGLSNARKFLVTSATNGDVPCSDPTGICQEDADRVSONSLGEMTK 892
 Db 1073 FLGPVIVFIPFALRG--KERELVYLSNGDSMKHEFCYTED-----ELNELIN 1122
 QY 893 LGSKTALNRDPTYRRQMOFTITROMSFDTKLVLTEDEWFSDISPQITWRLNLTIVSYGR 952
 Db 1123 -GMDEVLDSPEDLEKRICRITITRDF-----POYFAVVS--RIKQDSNLIGPEG 1169
 QY 953 LLRAN-----QISFWMDLASMINTEQWPRYTSWLLIYLETEGIPQOMLKTIVERIS 1007
 Db 1170 VLSSTVPQOVAPEFEGALTRIRY-----GLQAQPMHSELVKIL 1210
 QY 1008 KNIPPTKDEPLELEIDGIRNEFEVLSRTPVLVARDVKVLPCTVNLDPKIRELIADVR 1067
 Db 1211 GNKAT--FSPIVTLEPRRRKFKPTMTIPV-----PK----- 1241
 QY 1068 AAREGISIGLA--YPLPLHGGPPRAPSISQPPSVCSSTFNGPPAGVYSPQPHSSY 1125
 Db 1242 -ASSDVMLNGFGGDAFLRL-----LCSIT-----GGTTPAO----- 1272
 QY 1126 YSGMTGPQHPFYNRGSGPAPGPVYLLN-----SLNVDA-----VCEKIKQI----- 1166
 Db 1273 WEDITGT-----PLTFVNECVSTTNVSARFWLIDCRQIQESVTFASQVYR 1319
 QY 1167 EGIQDSMLPYCCTTIKKAN-INGRYLAQCNI--ELKEMNMNGDWHLEFRSTVLEMRN 1222
 Db 1320 EITCVYAKKFVFAKSHDPIEARLCFCMTDDKVDKTLBOQENPAE--VARSRDVEYLE 1377
 QY 1223 AESHYVPDPRLSSSSGPARPGPARASHNELP---HTELSQTPY-TLNSFPEELN 1278
 Db 1378 GKPIYVDFGMLVPLTKSG-QHHLFSFPAFKEENRLPLFKVVRDVTQEPGRLSPMKPEKS 1436
 QY 1279 TLGLDEGAPRHSNLSWOSOTRTPSLSLNSODSSIETSKLTDKVAQAEYRDAYREYIAQM 1338
 Db 1437 TRGLVHQAICNINILPTITKSES-----DOBOEEIDMTSEKNPDQERIEERLAYI 1491
 QY 1339 SQLEGPGSTTISGR-----SSPHSTYYMGSSSGSIHSLNLEQKGDSEPPKDDGRK 1392
 Db 1492 ADHLG--FSWTELARELDFTEBQIHQRIENPNSLOQSHALIKYLERDGHATDTNLV 1549
 QY 1393 SFILM--RGVDYIDSSSVSTNDASPLDPTIEDEKSDQSGSKLLPGKKSERSSLFOTD 1450
 Db 1550 ECLTQIRHMDIVHL---METN---TEPLQERISHSYAEIQTIT-LDHSEGFSLQEE 1600
 QY 1451 LKLKSGGLRYOKLPDEDESGTEESDNTPLKD-----DKDRKAGKVERVPKS 1499
 Db 1601 LCTAOKHOKEREAVKESSET-----CDHPPIVSEEDISVGYSTFQDGVKPTGDDSSSTALF 1656
 QY 1500 PEHSAEPIRTFIKAK-----EYLSDALLDKDDSDSGVRSSE-----SSPNH 1541
 Db 1657 PQTHEQYVOQDFSGKQDLPRESSLEYQOEYFV---TTPGTETSETQKAMIVSSPSK 1711
 QY 1542 SLHNFVADDSQLEKANLIELEDDSHSGKRGIPHSLSGLQDPIIARMSICSE--DKKSPSE 1599
 Db 1712 T-PEEVSTPAEERK---LYLOTPTSSEBGG-----SPLOEPPEPSSEHRESSPRK 1758

QY 1600 CSLIASSPEENMP-----ACOKAYNLNR-TPSTVTLNNSAPARANQNFDEMG 1648
 Db 1759 TSLVYESADNDPETERLDEDAAFKGDMPFIEPIYV-----EBEY 1802
 QY 1649 IRE-----TSOYLIRPSSS-----PNPTTIONENIKSMTHKRSQSS 1685
 Db 1803 IDEHGHVYKVTYRKIIIRRYVSESETEKEIMVQMPQEPVNIIEGQGYSVIKRVLKS 1862
 QY 1686 YTRLKSDPPE 1695
 Db 1863 DFEQSEBDNNE 1872

Search completed: July 1, 2003, 14:37:02
 Job time : 61.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2003, 14:34:04 ; Search time 89 Seconds
(without alignments)
3970.458 Million cell updates/sec

Title: US-10-021-571-2

Sequence: 1 MSYLISQSVINVEENIPALKEKCDVDERNECGTPTMLAAEGGNVEIVKELKN 1715

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

arched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.podent:*
12: sp.virus:*
13: sp.vertibrate:*
14: sp.unclassified:*
15: sp.virus:*
16: sp.bacteriap:*
17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	8853	100.0	1715 11 09ERD4	09erd4 rattus norv
2	8639.5	97.6	1762 11 09ERG6	09erg6 rattus norv
3	8212	92.8	1777 4 09ULH0	09ulh0 homo sapien
4	5073.5	57.3	1031 4 09U42	09u42 homo sapien
5	2408	27.2	543 4 09H9E4	09h9e4 homo sapien
6	2376.5	26.8	1498 5 09W211	09w211 drosophila
7	1848.5	20.9	1398 5 020109	020109 caenorhabdi
8	1024	11.6	246 4 09NT37	09nt37 homo sapien
9	599.5	6.8	129 4 09H889	09h889 homo sapien
10	555	6.3	426 16 08YTG9	08ytg9 anabaena sp
11	555	6.3	1856 4 099407	099407 homo sapien
12	551	6.2	1943 11 061307	061307 mus musculu
13	545	6.2	1719 4 013768	013768 homo sapien
14	529	6.0	2622 11 070511	070511 rattus norv
15	526	5.9	4377 4 012955	012955 homo sapien
16	516.5	5.8	1848 11 061302	061302 mus musculu

17	510	5.8	1726 11 08VC68	08vc68 mus musculu
18	505.5	5.7	1136 6 09N180	09n180 bos taurus
19	504	5.7	2443 5 09VSA2	09vsa2 drosophila
20	503.5	5.7	843 11 P97582	P97582 rattus norv
21	500.5	5.7	1486 4 08TEF1	08tef1 homo sapien
22	499.5	5.6	1159 5 09NCP8	09ncp8 drosophila
23	493.5	5.6	833 4 096186	096186 homo sapien
24	492.5	5.6	784 5 09VCAT	09vc47 drosophila
25	492.5	5.6	4001 5 08WR07	08wr07 drosophila
26	490.5	5.5	6994 5 017343	017343 caenorhabdi
27	488.5	5.5	1009 5 08SWY2	08swy2 drosophila
28	486.5	5.5	1599 11 09NHD0	09nhd0 mus musculu
29	483.5	5.5	1762 11 088521	088521 rattus norv
30	479.5	5.4	1188 4 09H288	09h288 homo sapien
31	479	5.4	2039 5 017489	017489 caenorhabdi
32	471	5.3	1282 5 08T4F8	08t4f8 drosophila
33	471	5.3	2119 5 09VAU5	09vaus drosophila
34	471	5.3	2119 5 08Y9G9	08y9g9 drosophila
35	467.5	5.3	1549 5 09V4B1	09v4b1 drosophila
36	464.5	5.2	1549 5 024241	024241 drosophila
37	459	5.2	743 4 08TB46	08tb46 homo sapien
38	455	5.1	460 4 09H6J9	09h6j9 homo sapien
39	454	5.1	1867 5 017486	017486 caenorhabdi
40	453	5.1	627 4 096G77	096g77 homo sapien
41	452.5	5.1	1809 5 017487	017487 caenorhabdi
42	451	5.1	1349 11 08VDM6	08vdm6 mus musculu
43	449.5	5.1	1088 4 013484	013484 homo sapien
44	443.5	5.0	1815 5 017488	017488 caenorhabdi
45	441	5.0	435 4 09HA95	09ha95 homo sapien

ALIGNMENTS

RESULT 1

ID 09ERD4 PRELIMINARY: PRT: 1715 AA.

AC 09ERD4: 01-MAR-2001 (TREMBLrel. 16, Created)

DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Ankyrin repeat-rich membrane-spanning protein.

GN ARMS.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.

OX NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=SPRAGUE-DAWLEY;

RX MEDLINE=20585245; PubMed=1150334;

RA Kong H., Boulter J., Weber J.L., Lai C., Chao M.V.;

RT "An evolutionarily conserved transmembrane protein that is a novel

RT downstream target of neurotrophin and ephrin receptors.";

RL J. Neurosci. 21:176-185(2001).

DR EMBL: AF313464; AAC34167.1; -.

DR HSSP: P80144; ZMYO.

DR InterPro: IPR002110; ANK.

DR Pfam: PF000023; ank_12.

DR PRINTS: PR01415; ANKYRIN.

DR SMART: SM00248; ANK_11.

DR PROSITE: PS50088; ANK_REPEAT; 10.

DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.

KW ANK repeat; Repeat.

SQ SEQUENCE 1715 AA; 190532 MW; 505593CA19A6CDC CRC64;

Qy 1 MSYLISQSVINVEENIPALKEKCDVDERNECGTPTMLAAEGGNVEIVKELKN 60

DB 1 MSYLISQSVINVEENIPALKEKCDVDERNECGTPTMLAAEGGNVEIVKELKN 60

Best Match 100.0%; Score 8853; DB 11; Length 1715;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 GANCLIEDNDNTALISASKEGHIIIVEELKSGASLEHRDMGWTALMAACVGRPDV 120
 DB 61 GANCLIEDNDNTALISASKEGHIIIVEELKSGASLEHRDMGWTALMAACVGRPDV 120
 QY 121 ELLLSHGAMPVTLQYSVPLIIMAGRGHADIYELLONGAKVCSDKYGTTPVMAAR 180
 DB 121 ELLLSHGAMPVTLQYSVPLIIMAGRGHADIYELLONGAKVCSDKYGTTPVMAAR 180
 QY 181 KGHLECVKHLAMGADVQEGANSMTALIVAKGCTOSVKEILLKRNPNVLLDKGNTA 240
 DB 181 KGHLECVKHLAMGADVQEGANSMTALIVAKGCTOSVKEILLKRNPNVLLDKGNTA 240
 QY 241 LMIAKEGHEIEVODLDAGTYVNIIPDRSGPTVLIGAVRGHVEIVRALLOKYADIDIRG 300
 DB 241 LMIAKEGHEIEVODLDAGTYVNIIPDRSGPTVLIGAVRGHVEIVRALLOKYADIDIRG 300
 QY 301 QDNKATLYAWAEKGNATVRDILQCNPDTEICTKGETPLIKATMRNIEVEVELLDKGA 360
 DB 301 QDNKATLYAWAEKGNATVRDILQCNPDTEICTKGETPLIKATMRNIEVEVELLDKGA 360
 QY 361 KVSAYDKGDPPLHVAIRGRSRRLAELLRNPKOSRLLYRNPKEGEPYNTIDCSHOKSIL 420
 DB 361 KVSAYDKGDPPLHVAIRGRSRRLAELLRNPKOSRLLYRNPKEGEPYNTIDCSHOKSIL 420
 QY 421 TOIFGARHLSPETDGDMLGYDLYSSALADILSEPTMOPICVGLYAOVSGSKSFLKL 480
 DB 421 TOIFGARHLSPETDGDMLGYDLYSSALADILSEPTMOPICVGLYAOVSGSKSFLKL 480
 QY 481 EDEMTFAGQOETPLFOFSMLIVPLTLLCGGLGVFPVVDNLAIAISLFLALYIF 540
 DB 481 EDEMTFAGQOETPLFOFSMLIVPLTLLCGGLGVFPVVDNLAIAISLFLALYIF 540
 QY 541 FVIFYGGRGREGESNNMAMALSTRARHIGYLELLEFKLMFNPPELPOTKALPVPLF 600
 DB 541 FVIFYGGRGREGESNNMAMALSTRARHIGYLELLEFKLMFNPPELPOTKALPVPLF 600
 QY 601 TDYNRLSSVGETSLAEMIATLSDACEREFGLATRLFRVTRTEESOGKKKKTKCPLPS 660
 DB 601 TDYNRLSSVGETSLAEMIATLSDACEREFGLATRLFRVTRTEESOGKKKKTKCPLPS 660
 QY 661 FVIFYGCIAGITLLAIFRVDPKHLTVNAIILISAVGLAFVLCNRTMOWYDLSL 720
 DB 661 FVIFYGCIAGITLLAIFRVDPKHLTVNAIILISAVGLAFVLCNRTMOWYDLSL 720
 QY 721 NSORRRLHSAASKLHLKLESEGFMVYLKCEVELMARMAKTIDSFQONOTRLVYIIDGDAC 780
 DB 721 NSORRRLHSAASKLHLKLESEGFMVYLKCEVELMARMAKTIDSFQONOTRLVYIIDGDAC 780
 QY 781 EODKVLQMLDLYRVLYFSKGPPIAIPASDPHIIKAINONLNSVLSDSNINSHDMYRNVH 840
 DB 781 EODKVLQMLDLYRVLYFSKGPPIAIPASDPHIIKAINONLNSVLSDSNINSHDMYRNVH 840
 QY 841 LPVFLNSRGLSNARKFLVTSATNDITCSDTTQEDTDKRVSONSISGEMTKLGSKALN 900
 DB 841 LPVFLNSRGLSNARKFLVTSATNDITCSDTTQEDTDKRVSONSISGEMTKLGSKALN 900
 QY 901 RRDYRRRQOMORTTROMSFDLTCLVTEWFSDISPQJMRRLINITSVGRLLRANOIT 960
 DB 901 RRDYRRRQOMORTTROMSFDLTCLVTEWFSDISPQJMRRLINITSVGRLLRANOIT 960
 QY 961 FNNRRLASWIMLTQOWPRTSWMLLYLEETEGLDQMTLKTMYERISKNIPTTKDVEPLL 1020
 DB 961 FNNRRLASWIMLTQOWPRTSWMLLYLEETEGLDQMTLKTMYERISKNIPTTKDVEPLL 1020
 QY 1021 EIDDIDIRNEVEFLSSRTPVLAARDVKTFLPCTVWLDKLEILLADVAARAOJNIGSLAY 1080
 DB 1021 EIDDIDIRNEVEFLSSRTPVLAARDVKTFLPCTVWLDKLEILLADVAARAOJNIGSLAY 1080
 QY 1081 PPLPLHRCPPRPSPGSGOPASVCSASFNGPFGGVVSPPHSSYSGLSGPOHPFNRA 1140
 DB 1081 PPLPLHRCPPRPSPGSGOPASVCSASFNGPFGGVVSPPHSSYSGLSGPOHPFNRA 1140

QY 1141 AVPATSSLLSSMTVDVYVCEKRLQTEGLDONMMPOYCTTIKANINGRVLSCNTDELK 1200
 DB 1141 AVPATSSLLSSMTVDVYVCEKRLQTEGLDONMMPOYCTTIKANINGRVLSCNTDELK 1200
 QY 1201 KEMAMNGDWHLEFRSWMLEKRSVESQVPEDEPFLLENDSAPVPHGESARRSRHTLEPLT 1260
 DB 1201 KEMAMNGDWHLEFRSWMLEKRSVESQVPEDEPFLLENDSAPVPHGESARRSRHTLEPLT 1260
 QY 1261 ELSQCPYTLNFSPEBELNTLGLDEGAPRHSNLSWSQTRTPSLSSLSQDSISIEISKLT 1320
 DB 1261 ELSQCPYTLNFSPEBELNTLGLDEGAPRHSNLSWSQTRTPSLSSLSQDSISIEISKLT 1320
 QY 1321 DKVOAERADYREYIAQMSQLEGTSSTISGSSPHSYTYIGQSSSGSISTLEQERG 1380
 DB 1321 DKVOAERADYREYIAQMSQLEGTSSTISGSSPHSYTYIGQSSSGSISTLEQERG 1380
 QY 1381 KEGELKQEDGRKSFLLMRGVDYIDSSSGVSTNEASPLDPTTEDEKSDQSGSKLLPGKKS 1440
 DB 1381 KEGELKQEDGRKSFLLMRGVDYIDSSSGVSTNEASPLDPTTEDEKSDQSGSKLLPGKKS 1440
 QY 1441 SERPSLEFQNDLKLKGGGLRYQKLPSPDEDSGTRVQITPHCSKMITRKLAKKORPCASP 1500
 DB 1441 SERPSLEFQNDLKLKGGGLRYQKLPSPDEDSGTRVQITPHCSKMITRKLAKKORPCASP 1500
 QY 1501 QESHAEPITRTIFAKKEYLSDALLDKRSDSGVRSNESSPNHSLHNEADDSQLKCANLI 1560
 DB 1501 QESHAEPITRTIFAKKEYLSDALLDKRSDSGVRSNESSPNHSLHNEADDSQLKCANLI 1560
 QY 1561 ELEDGHSGRKGMPSHLSGLDPTIARMSTICSEDKKSPESCSLIASSPESWPACOKAVN 1620
 DB 1561 ELEDGHSGRKGMPSHLSGLDPTIARMSTICSEDKKSPESCSLIASSPESWPACOKAVN 1620
 QY 1621 LNRTPSTVTLLNNATPTNANONFDEIGRETSOYLILPSPNPTAVONNLSKMAH 1680
 DB 1621 LNRTPSTVTLLNNATPTNANONFDEIGRETSOYLILPSPNPTAVONNLSKMAH 1680
 QY 1681 RSQRSSYTRLKSDASELHAASSESTGFGERSIL 1715
 DB 1681 RSQRSSYTRLKSDASELHAASSESTGFGERSIL 1715
 RESULT 2
 Q9EOG6
 ID Q9EOG6 PRELIMINARY; PRT; 1762 AA.
 AC Q9EOG6;
 DT 01-MAR-2001 (T-EMBLrel. 16, Created)
 DT 01-OCT-2001 (T-EMBLrel. 18, Last sequence update)
 DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
 DE KIDINS220.
 GN KIDINS220.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NBL_TaxID-101116;
 RN [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE-20568256; PubMed-10998417;
 RA Iglesias T., Cabrera-Poch N., Mitchell M.P., Naven T.J., Rozenfurt E.,
 RA Schiavo G.;
 RT "Identification and cloning of Kidins220, a novel neuronal substrate
 of protein kinase D.";
 RL J. Biol. Chem. 275:40048-40056(2000).
 RN [2]
 RN SEQUENCE FROM N.A.
 RA Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RN SEQUENCE FROM N.A.
 RA Mitchell M., Schiavo G.;
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AR33045; ANG35185.2;
 DR HSSP: P80144; 2MYO.
 DR InterPro: IPR002110; ANK.

DR pfam: PF00023; ank: 12.
 DR PRINTS: PRO1415; ANKYRIN.
 DR SMART: SM00248; ANK: 11.
 DR PROSITE: PS0088; ANK_REPEAT; 10.
 DR ANK repeat; Repeat.
 SW SOURCE 1762 AA; 195715 MW; OCB2689A57JF8AB4 CRC64;

Query Match 97.6%; Score 8639.5; DB 11; Length 1762;
 Best Local Similarity 95.8%; Pred. No. 0;
 Matches 1691; Conservative 6; Mismatches 15; Indels 53; Gaps 5;

QY 1 MSVLSQSYINVEEENIPALKALLKCKDVDERNECGOTPLMLAEOGNVEIVKELKN 60
 DB 1 MSVLSQSYINVEEENIPALKALLKCKDVDERNECGOTPLMLAEOGNVEIVKELKN 60
 QY 61 GANCNEDDDNNTALLISASKEGHIHYEELKSGASLEHRDGGWTALMACYKGGTDVY 120
 DB 61 GANCNEDDDNNTALLISASKEGHIHYEELKSGASLEHRDGGWTALMACYKGGTDVY 120
 QY 121 ELLSHGAMPSTYGLQYVYPIYMAARGHADIYHLLONGAKVNCSDKYGTTPYMAAR 180
 DB 121 ELLSHGAMPSTYGLQYVYPIYMAARGHADIYHLLONGAKVNCSDKYGTTPYMAAR 180
 QY 181 KGHLECVKHLANGADVDEGANSMTALIVAAKGYSQVKEILKRNPNVNLTDKXGNTA 240
 DB 180 KGHLECVKHLANGADVDEGANSMTALIVAAKGYSQVKEILKRNPNVNLTDKXGNTA 239
 QY 241 LMAISKEGHEIYODLLDAGTYNIPDRSGDTYLIGAVRGHVEIYRALLQKADIDING 300
 DB 240 LMAISKEGHEIYODLLDAGTYNIPDRSGDTYLIGAVRGHVEIYRALLQKADIDING 299
 QY 301 QDNKTLAYMAVEKGNATWVDILQCNPDTEICTKDETFPLIKATKRNIEVEELLDKGA 360
 DB 300 QDNKTLAYMAVEKGNATWVDILQCNPDTEICTKDETFPLIKATKRNIEVEELLDKGA 359
 QY 361 KYSADVKKGDTPLHVAIRGSRRLAELLRNPKDGLLYRNKAGETPYNIDCSHOKSIL 420
 DB 360 KYSADVKKGDTPLHVAIRGSRRLAELLRNPKDGLLYRNKAGETPYNIDCSHOKSIL 419
 QY 421 TOIFGARHLSPTEDMDLGYDLYSSALADILSEPMOPICVGLAOMSGSFLKTL 480
 DB 420 TOIFGARHLSPTEDMDLGYDLYSSALADILSEPMOPICVGLAOMSGSFLKTL 479
 QY 481 EDEMKTFAGQOTEPLEFQSWLIVETLLCGGLGFAFVPVDNTLAIALISFLALITYF 540
 DB 480 EDEMKTFAGQOTEPLEFQSWLIVETLLCGGLGFAFVPVDNTLAIALISFLALITYF 539
 QY 541 FIYIYGGRRGESMMAMALSTRLARHIGYLELLEKMFVNPELPEOTYKALPVRLF 600
 DB 540 FIYIYGGRRGESMMAMALSTRLARHIGYLELLEKMFVNPELPEOTYKALPVRLF 599
 QY 601 TDNRLSSVGGETSLAEMTATLSDACEREGELATLFRFRREESOGKKKKKCTCLTS 660
 DB 600 TDNRLSSVGGETSLAEMTATLSDACEREGELATLFRFRREESOGKKKKKCTCLTS 659
 QY 661 FVLEFLVGCIIAGITLLAFRVDPKHLTYNALISIASVYGLAFNLRTMWOVDSIL 720
 DB 660 FVLEFLVGCIIAGITLLAFRVDPKHLTYNALISIASVYGLAFNLRTMWOVDSIL 719
 QY 721 NSGRKLHSAASKLHKLSGEPKVLKCEVELMARNAKTIIDSTQOTRLVITIGLDAC 780
 DB 720 NSGRKLHSAASKLHKLSGEPKVLKCEVELMARNAKTIIDSTQOTRLVITIGLDAC 779
 QY 781 EOKKVLQMDLTVAVLSKGPFIAPASDPHIIKAIKONINSVLRKSNINGHYMNITYH 840
 DB 780 EOKKVLQMDLTVAVLSKGPFIAPASDPHIIKAIKONINSVLRKSNINGHYMNITYH 839
 QY 841 LPVFLNSRGLSNAKRLVYATNGDITCSDTTGTQEDTDRYSONSLGEMTKLGSXTALN 900
 DB 840 LPVFLNSRGLSNAKRLVYATNGDITCSDTTGTQEDTDRYSONSLGEMTKLGSXTALN 899
 QY 901 RROTYYRRQOMQRTITROMSFDLKLVTEDWESDISQYRRLNLIVSYTGRLLRANOIT 960
 DB 901 RROTYYRRQOMQRTITROMSFDLKLVTEDWESDISQYRRLNLIVSYTGRLLRANOIT 960

DB 900 RROTYYRRQOMQRTITROMSFDLKLVTEDWESDISQYRRLNLIVSYTGRLLRANOIT 959
 QY 961 FNMWDLASWINTEQWPRRTSMILLYLEETGGLDQWMLKMYERISNITPTKDVPEL 1020
 DB 960 FNMWDLASWINTEQWPRRTSMILLYLEETGGLDQWMLKMYERISNITPTKDVPEL 1019
 QY 1021 EIDGIRNFEVFLSRTPEVLVARDVKTPLPCTVNDLPKRELIADVRAARQINIGLAY 1080
 DB 1020 EIDGIRNFEVFLSRTPEVLVARDVKTPLPCTVNDLPKRELIADVRAARQINIGLAY 1079
 QY 1081 PPLPLEGPPRPPSGYSQPASVCSASFNGPPGGVSPQPHSSYSGLSGPPHFPYNR 1139
 DB 1080 PPLPLEGPPRPPSGYSQPASVCSASFNGPPGGVSPQPHSSYSGLSGPPHFPYNR 1139
 QY 1140 -----AAPTATGSSILL 1151
 DB 1140 -----AAPTATGSSILL 1151
 QY 1152 SSMVDVVCERLQJTEGLDQNMMPQCYTIKKANINGVLSQCNIDELKEMANFGDMH 1211
 DB 1200 SSMVDVVCERLQJTEGLDQNMMPQCYTIKKANINGVLSQCNIDELKEMANFGDMH 1209
 QY 1212 LFRSVLEMRVYESOVVEDPPLNENSSAPVPHGESARRSHHELPLELSSOTPYTLN 1271
 DB 1260 LFRSVLEMRVYESOVVEDPPLNENSSAPVPHGESARRSHHELPLELSSOTPYTLN 1319
 QY 1272 PSEELNLTGLDEGAPRHSNLSWOSOTRRTSLNSODSISTEISKTLDVQAEYRAY 1331
 DB 1320 PSEELNLTGLDEGAPRHSNLSWOSOTRRTSLNSODSISTEISKTLDVQAEYRAY 1379
 QY 1332 REYIAQMSOLEGGTSGSTISGRSPHSTYIYGOSSSGGSISHTLEOERKGEELKOEDEGR 1391
 DB 1380 REYIAQMSOLEGGTSGSTISGRSPHSTYIYGOSSSGGSISHTLEOERKGEELKOEDEGR 1439
 QY 1392 KSFLLKRGVDYIDSSSGYSTNEASPLDYTEDEKSDOSGSKLPGKSSRPSLFTDL 1451
 DB 1440 KSFLLKRGVDYIDSSSGYSTNEASPLDYTEDEKSDOSGSKLPGKSSRPSLFTDL 1499
 QY 1452 KLGKGLRYOKLPSEDESGTGRVQITPHCSMTITRKLLAK-ORECASPOEHSAPERT 1510
 DB 1500 KLGKGLRYOKLPSEDESGTGRVQITPHCSMTITRKLLAK-ORECASPOEHSAPERT 1557
 QY 1511 FIKAKELYSLDALDKDSDSGVRSNESPNSHLENAADDSOLEKANILEDEGSHGK 1570
 DB 1558 FIKAKELYSLDALDKDSDSGVRSNESPNSHLENAADDSOLEKANILEDEGSHGK 1617
 QY 1571 RGMPSLSGLDPTIARMSICSEDKSPSEGLIASPEESNPACQKAYNLNRTPSYTL 1630
 DB 1618 RGMPSLSGLDPTIARMSICSEDKSPSEGLIASPEESNPACQKAYNLNRTPSYTL 1677
 QY 1631 NNTAPPTANONFPEIGIRETSOVLIRPGSPMPRTAVONENKSMHKKRSOSSYTRL 1690
 DB 1678 NNTAPPTANONFPEIGIRETSOVLIRPGSPMPRTAVONENKSMHKKRSOSSYTRL 1737
 QY 1691 SKDASELHAASSESTGFGEERSIL 1715
 DB 1738 SKDASELHAASSESTGFGEERSIL 1762

RESULT 3
 Q9ULH0 PRELIMINARY; PRT: 1777 AA.
 AC Q9ULH0;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-OCT-2001 (TREMblrel. 18, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE KIAA1250 protein (Fragment).
 GN KIAA1250.
 OS Homo sapiens (Human).
 OC Eukaryota; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]

RP SEQUENCE FROM N.A.
 RC TISSUE-BRAIN;
 RX MEDLINE-20039619; PubMed-10574462;
 RA Nagase T., Ishikawa K., Kikuno R., Hirosewa M., Nomura N., Ohara O.,
 RT "Prediction of the coding sequences of unidentified human genes. XV.
 RT The complete sequences of 100 new cDNA clones from brain which code
 RT for large proteins in vitro."
 RL DNA Res. 6:337-345(1999).
 DR EMBL: AB033076; BAA86564.2; -
 DR Interpro: IPR02110; ANK.
 DR Pfam: PF00023; ank, 12.
 DR PRINTS: PRO1415; ANKTRIN.
 DR SMART: SM00248; ANK, 11.
 DR PROSITE: PS50088; ANK_REPEAT, 10.
 DR PROSITE: PS50297; ANK_REPEAT, 1.
 DR ANK repeat; Repeat.
 FT NON_TER
 SQ SEQUENCE 1777 AA; 197209 MW; B6505923FB45F143 CRC64;

Query Match 92.8%; Score 8212; DB 4; Length 1777;
 Best Local Similarity 89.7%; Pred. No. 0;

Matches 1590; Conservative 53; Mismatches 70; Indels 60; Gaps 4;

DB 1 MSVLISQSYINVEENIPALKKLECKDVERNECGQTPMLAEQGNVEIYKELIKN 60
 QY 61 GANCNLEPDMNTALISASKGHIHVEELKSGASLEHRDMGCTALMAACYGRDYY 120
 DB 67 GANCNLEPDMNTALISASKGHIHVEELKSGASLEHRDMGCTALMAACYGRDYY 126
 QY 121 ELLSHGANPSVTGLQYVYPTIWAAGRHADIVALLQNGAKVNSDKYGTTPWMAAR 180
 DB 127 ELLSHGANPSVTGLQYVYPTIWAAGRHADIVALLQNGAKVNSDKYGTTPWMAAR 185
 QY 181 KGHLECYKHLAMGADVQEGANSKTALIVAKGTYTSVKEILKRNPNVLTJDKGNTA 240
 DB 186 KGHLECYKHLAMGADVQEGANSKTALIVAKGTYTSVKEILKRNPNVLTJDKGNTA 245
 QY 241 LMAISKEGHIETVDLDAGTYVNPDRSGPVLIGAVRGHVEIVRALLOKYADIDIRG 300
 DB 246 LMAISKEGHIETVDLDAGTYVNPDRSGPVLIGAVRGHVEIVRALLOKYADIDIRG 305
 QY 301 QDNKTALYWAWEKGNATWVRILOCPNPTETCTKDGTEPLIKATKRNIEVEVELLDKA 360
 DB 306 QDNKTALYWAWEKGNATWVRILOCPNPTETCTKDGTEPLIKATKRNIEVEVELLDKA 365
 QY 361 KVSANDKGDPTPLHAIIGRSRRLAELLRNPKGRLLYRNKAGETPTYNIDCSHQSIL 420
 DB 366 KVSANDKGDPTPLHAIIGRSRRLAELLRNPKGRLLYRNKAGETPTYNIDCSHQSIL 425
 QY 421 TQIFGARHLSPTETDGMIGYDIYSSALADILSEPTMOPICVGLYAOMSGKSFLLKL 480
 DB 426 TQIFGARHLSPTETDGMIGYDIYSSALADILSEPTMOPICVGLYAOMSGKSFLLKL 485
 QY 481 EDEKTPFGQOTEPPLQFWSLIVFTLLCGSLGVFAEPAVDNLAIAISLSTALIYIF 540
 DB 486 EDEKTPFGQOTEPPLQFWSLIVFTLLCGSLGVFAEPAVDNLAIAISLSTALIYIF 545
 QY 541 FIVIFGRRGREGSNMAMALSTRARHIGTLELLFKMFVNPELBPOTKALPVRFLE 600
 DB 546 FIVIFGRRGREGSNMAMALSTRARHIGTLELLFKMFVNPELBPOTKALPVRFLE 605
 QY 601 TDYNTLSVSGETSLAEIATLSDACEREFGLATRLFRVPTESOGKKKKTKCCIPS 660
 DB 606 TDYNTLSVSGETSLAEIATLSDACEREFGLATRLFRVPTESOGKKKKTKCCIPS 665
 QY 661 FVIFLPIVGCITAGITLLAIFRVDPKHLTVNAILISIASVGLAFVINCRTMVOVDSLL 720
 DB 666 FVIFLPIVGCITAGITLLAIFRVDPKHLTVNAILISIASVGLAFVINCRTMVOVDSLL 725
 QY 721 NSQRRLHNSASKLHKLSSEGFMKVLCVEELMARMAKTIDSTFQNOTRLVITIDGDAC 780

DB 726 NSQRRLHNSASKLHKLSSEGFMKVLCVEELMARMAKTIDSTFQNOTRLVITIDGDAC 785
 QY 781 EOKKVLQMDTVAVLUSKGPFIAPASDPHIIIIKAINONLSVLRBSNINGDHMYNIYH 840
 DB 786 EOKKVLQMDTVAVLUSKGPFIAPASDPHIIIIKAINONLSVLRBSNINGDHMYNIYH 845
 QY 841 LPVFLNSRGLSNARKFLVTSATNGDITCSPTTGTQEDTPRVSQNSLGEWTKLGSSTALN 900
 DB 846 LPVFLNSRGLSNARKFLVTSATNGDITCSPTTGTQEDTPRVSQNSLGEWTKLGSSTALN 905
 QY 901 RDTYRRRQMQRTTROMSFDLTKLVTEDEWFSIDISPTWRRLNIVSYTGRLRANOIT 960
 DB 906 RDTYRRRQMQRTTROMSFDLTKLVTEDEWFSIDISPTWRRLNIVSYTGRLRANOIT 965
 QY 961 FNMDRLASWINTLEOWMYRTSWILVLEETEGIPDQMTLKTWERISKNIPTTKDVEPLL 1020
 DB 966 FNMDRLASWINTLEOWMYRTSWILVLEETEGIPDQMTLKTWERISKNIPTTKDVEPLL 1025
 QY 1021 EIDGDIRNFEVFLSSRTPVLYARDVTFPLCTVNLDPKLEITADYRARAREQINIGGLAY 1080
 DB 1026 EIDGDIRNFEVFLSSRTPVLYARDVTFPLCTVNLDPKLEITADYRARAREQINIGGLAY 1085
 QY 1081 PPLPLEHGPFRPSPGTSQPASVCSASFNPGPPGVSPQPHSYISGLSGQHPHYN-- 1138
 DB 1086 PPLPLEHGPFRPSPGTSQPASVCSASFNPGPPGVSPQPHSYISGLSGQHPHYN-- 1145
 QY 1139 -----RAANP 1143
 DB 1146 FPAAPLYTPRYVGGSOHLISRPVKTSLPDRQNGLEVIKEDAAEGLSPPDSGSGSP 1205
 QY 1144 ATGSSLLSSMTYDVVCCERLQJEGHDOQMMPOYCTTIKANKINGVLSQCNIDELKCKM 1203
 DB 1206 APGVVVLNLANVDANVCERLQJEGHDOQMMPOYCTTIKANKINGVLSQCNIDELKCKM 1265
 QY 1204 AANFGWHLFRSVNLEKRSYVSEVYEDPAPFLNENSAAPVPGESARRSSHELPLETSL 1263
 DB 1266 AANFGWHLFRSVNLEKRSYVSEVYEDPAPFLNENSAAPVPGESARRSSHELPLETSL 1325
 QY 1264 SOTPYTLNFSFEELNTGLDEGAPRHSNLSWQOTRTPSLSSLNQDSISEISKLTIDY 1323
 DB 1326 SOTPYTLNFSFEELNTGLDEGAPRHSNLSWQOTRTPSLSSLNQDSISEISKLTIDY 1385
 QY 1324 QAEYRDAYREYIAQMSQLEGGTSTISGRSPHSTYIIGQSSGGSIHSTLEOERKXG 1383
 DB 1386 QAEYRDAYREYIAQMSQLEGGTSTISGRSPHSTYIIGQSSGGSIHSTLEOERKXG 1445
 QY 1384 ELKQEOGRKSFLLKRGDIVIDSSGYSTNAPSLDITFEDEKSDQSGSKLLPKKSSSR 1443
 DB 1446 ELKQEOGRKSFLLKRGDIVIDSSGYSTNAPSLDITFEDEKSDQSGSKLLPKKSSSR 1505
 QY 1444 PSLFQTDLKLKGGGLRYOKLPSEDESGCGRAVOITPHCSKMITRKLAKORECASPQH 1503
 DB 1506 PSLFQTDLKLKGGGLRYOKLPSEDESGCGRAVOITPHCSKMITRKLAKORECASPQH 1564
 QY 1504 SAEPITRTIAKEYISDALDKKSDSGVRSNESPNSHLNEAADSOLEKANLIELE 1563
 DB 1565 SAEPITRTIAKEYISDALDKKSDSGVRSNESPNSHLNEAADSOLEKANLIELE 1624
 QY 1564 DEHSGKRGMPHSLSLQDPTIARMSITCEDKSPSEGLIASSPEESPACOKAYNINR 1623
 DB 1625 DEHSGKRGMPHSLSLQDPTIARMSITCEDKSPSEGLIASSPEESPACOKAYNINR 1684
 QY 1624 TPSTVTLNNNTAPNTANANFDEIEIRTSOVIILPPGSPPTAVONENLSMAHKSQ 1683
 DB 1685 TPSTVTLNNNTAPNTANANFDEIEIRTSOVIILPPGSPPTAVONENLSMAHKSQ 1744
 QY 1684 RSSYTRLSKDBSELH--AASSESTGGEERESIL 1715
 DB 1745 RSSYTRLSKDBSELH--AASSESTGGEERESIL 1777

RESULT 4
 Q90F42

ID 09UP42 PRELIMINARY; PRT; 1031 AA.
 AC 09UP42;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Hypothetical 115.3 kDa protein.
 GN DKEFP434F0621.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=TESTIS;
 RA Blum H., Bauersachs S., Mewes H.W., Gassenhuber J., Wiemann S.;
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL133620; CAB63746.1;
 DR HSSP; P80144; 2MYO.
 DR InterPro: IPR002110; ANK.
 DR Pfam: PF000023; ank. 12.
 DR PRINTS; PR01415; ANKYRIN.
 DR SMART; SM00248; ANK; 11.
 DR PROSITE; PS50088; ANK_REPEAT; 10.
 DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
 DR ANK repeat: Hypothetical protein; Repeat.
 SO SEQUENCE 1031 AA; 11531 MW; D9E866461C13A2F5 CRC64;

Query Match 57.3%; Score 5073.5; DB 4; Length 1031;
 Best Local Similarity 95.2%; Pred. No. 0;
 Matches 974; Conservative 23; Mismatches 13; Indels 13; Gaps 1;

QY 1 MSVLTSSVINYEEENIPALKECKVDENECGQPLMAAOGVNEIVKELTKN 60
 DB 1 MSVLTSSVINYEEENIPALKECKVDENECGQPLMAAOGVNEIVKELTKN 60
 QY 61 GANCNEDLDNMTALISASKEGHIIVEELLESASLEHNDMGMTALMACYKGRDYY 120
 DB 61 GANCNEDLDNMTALISASKEGHIIVEELLESASLEHNDMGMTALMACYKGRDYY 120
 QY 61 GANCNEDLDNMTALISASKEGHIIVEELLESASLEHNDMGMTALMACYKGRDYY 120
 DB 61 GANCNEDLDNMTALISASKEGHIIVEELLESASLEHNDMGMTALMACYKGRDYY 120
 QY 121 ELLISHGANSVYGLQSYVPIIWAAGRHADIVHLLONGAVNCSDKGTTPVWAAR 180
 DB 121 ELLISHGANSVYGLQSYVPIIWAAGRHADIVHLLONGAVNCSDKGTTPVWAAR 180
 QY 121 ELLISHGANSVYGLQSYVPIIWAAGRHADIVHLLONGAVNCSDKGTTPVWAAR 180
 DB 121 ELLISHGANSVYGLQSYVPIIWAAGRHADIVHLLONGAVNCSDKGTTPVWAAR 180
 QY 181 KGHLECYKHLAMGADVDOGANSMTALIAVKGQSYQSKELIKRPNVNLDDKQNTA 240
 DB 181 KGHLECYKHLAMGADVDOGANSMTALIAVKGQSYQSKELIKRPNVNLDDKQNTA 240
 QY 181 KGHLECYKHLAMGADVDOGANSMTALIAVKGQSYQSKELIKRPNVNLDDKQNTA 240
 DB 181 KGHLECYKHLAMGADVDOGANSMTALIAVKGQSYQSKELIKRPNVNLDDKQNTA 240
 QY 241 LMISASKEGHIIVEELLESASLEHNDMGMTALMACYKGRDYY 300
 DB 241 LMISASKEGHIIVEELLESASLEHNDMGMTALMACYKGRDYY 300
 QY 241 LMISASKEGHIIVEELLESASLEHNDMGMTALMACYKGRDYY 300
 DB 241 LMISASKEGHIIVEELLESASLEHNDMGMTALMACYKGRDYY 300
 QY 301 QDNKTALYAVEKNATWVRDILQCNPDTEICTGDETPPLKATKMNIEVELLDKGA 360
 DB 301 QDNKTALYAVEKNATWVRDILQCNPDTEICTGDETPPLKATKMNIEVELLDKGA 360
 QY 301 QDNKTALYAVEKNATWVRDILQCNPDTEICTGDETPPLKATKMNIEVELLDKGA 360
 DB 301 QDNKTALYAVEKNATWVRDILQCNPDTEICTGDETPPLKATKMNIEVELLDKGA 360
 QY 361 KVSADVKKGDPPLHATRGSRRLAELLRNPKDGLLYRNKGEPPYNDCHOKSIL 420
 DB 361 KVSADVKKGDPPLHATRGSRRLAELLRNPKDGLLYRNKGEPPYNDCHOKSIL 420
 QY 361 KVSADVKKGDPPLHATRGSRRLAELLRNPKDGLLYRNKGEPPYNDCHOKSIL 420
 DB 361 KVSADVKKGDPPLHATRGSRRLAELLRNPKDGLLYRNKGEPPYNDCHOKSIL 420
 QY 421 TQIFGARHLSPTETDGMGLGYDLSALADILSEPTQOPICVGLYAQWGSKSFLLKL 480
 DB 421 TQIFGARHLSPTETDGMGLGYDLSALADILSEPTQOPICVGLYAQWGSKSFLLKL 480
 QY 421 TQIFGARHLSPTETDGMGLGYDLSALADILSEPTQOPICVGLYAQWGSKSFLLKL 480
 DB 421 TQIFGARHLSPTETDGMGLGYDLSALADILSEPTQOPICVGLYAQWGSKSFLLKL 480
 QY 481 EDEKMTFAGQOTEPPLFQPSWILVLTLLCGGLVNAFPVDTMLAIAISFLALYIF 540
 DB 481 EDEKMTFAGQOTEPPLFQPSWILVLTLLCGGLVNAFPVDTMLAIAISFLALYIF 540
 QY 481 EDEKMTFAGQOTEPPLFQPSWILVLTLLCGGLVNAFPVDTMLAIAISFLALYIF 540
 DB 481 EDEKMTFAGQOTEPPLFQPSWILVLTLLCGGLVNAFPVDTMLAIAISFLALYIF 540
 QY 541 FIVITFGGRGEGSNMAMALSTRARHIGVLELFLKMFVNPELPEOTKALPVLEIF 600
 DB 541 FIVITFGGRGEGSNMAMALSTRARHIGVLELFLKMFVNPELPEOTKALPVLEIF 600
 QY 541 FIVITFGGRGEGSNMAMALSTRARHIGVLELFLKMFVNPELPEOTKALPVLEIF 600
 DB 541 FIVITFGGRGEGSNMAMALSTRARHIGVLELFLKMFVNPELPEOTKALPVLEIF 600
 QY 601 TDYNNLSSVGETSLAEMIATLSDACEREFGLATRLFRVTRTEESOGKKKKKTCCLPS 660
 DB 601 TDYNNLSSVGETSLAEMIATLSDACEREFGLATRLFRVTRTEESOGKKKKKTCCLPS 660

DB 601 TDYNNLSSVGETSLAEMIATLSDACEREFGLATRLFRVTRTEESOGKKKKKTCCLPS 660
 QY 661 FVIFPLFYGCIAGTTLAIPRVDPKHLTVANAIIISTASVYGLAFVNLCTRMQVLDL 720
 DB 661 FVIFPLFYGCIAGTTLAIPRVDPKHLTVANAIIISTASVYGLAFVNLCTRMQVLDL 720
 QY 721 NSQRRRLHNAASKLKLKSEGMKYLKCEVELMAMAKTISFTQNTRLVVIIDGLDAC 780
 DB 721 NSQRRRLHNAASKLKLKSEGMKYLKCEVELMAMAKTISFTQNTRLVVIIDGLDAC 780
 QY 781 EODKVLQMDIVRVLESKGPFLAIPASDPHIIKAIINONLSVLDNSINCHDYMRIVH 840
 DB 781 EODKVLQMDIVRVLESKGPFLAIPASDPHIIKAIINONLSVLDNSINCHDYMRIVH 840
 QY 841 LPVPLNSRLSNARKFVLTATNGDITGSDPTGQEDPDNRSONSGEMTKLSKALN 900
 DB 841 LPVPLNSRLSNARKFVLTATNGDITGSDPTGQEDPDNRSONSGEMTKLSKALN 900
 QY 901 RRDYRRROMQRTIRONSFDLTKLVTEEDWFSDISPQTMRLINIVSVTGLLRANOIT 960
 DB 901 RRDYRRROMQRTIRONSFDLTKLVTEEDWFSDISPQTMRLINIVSVTGLLRANOIT 960
 QY 961 FVMDRLASWINTEQAPRTSWLLIYEETBGLPDQMTLKTYER-----IS 1007
 DB 961 FVMDRLASWINTEQAPRTSWLLIYEETBGLPDQMTLKTYER-----IS 1007
 QY 1008 KNI 1010
 DB 1021 KSV 1023

RESULT 5

ID 09H9E4 PRELIMINARY; PRT; 543 AA.
 AC 09H9E4;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)
 DE CDNA FLJ12811 fis, clone NT2RP2002475.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
 RA Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,
 RA Nakamura Y., Nagahari K., Masuko Y., Sasaki N.;
 RT "MEDO human cDNA sequencing project."
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK022873; BAB14285.1; -
 SO SEQUENCE 543 AA; 55976 MW; F5D643D5A20C641D CRC64;

Query Match 27.2%; Score 2408; DB 4; Length 543;
 Best Local Similarity 86.8%; Pred. No. 1,3e-148;
 Matches 472; Conservative 24; Mismatches 46; Indels 2; Gaps 2;

QY 1173 MHPQCTTKKANINGRVLSQCNIDELKEMANPCDMLHFRSNVLEMSVSQVPPEDP 1232
 DB 1 MHPQCTTKKANINGRVLSQCNIDELKEMANPCDMLHFRSNVLEMSVSQVPPEDP 1232
 QY 1233 RFLNENSSAPVPGESARSSSHLELPUTELSSOTPYTLNFSPEELTGLDGCAPRSHNL 1292
 DB 1233 RFLNENSSAPVPGESARSSSHLELPUTELSSOTPYTLNFSPEELTGLDGCAPRSHNL 1292
 QY 61 RFLNENSSAPVPGESARSSSHLELPUTELSSOTPYTLNFSPEELTGLDGCAPRSHNL 120
 DB 61 RFLNENSSAPVPGESARSSSHLELPUTELSSOTPYTLNFSPEELTGLDGCAPRSHNL 120
 QY 1293 SMOQSTRPPLSLNSQDSISIKLTKVQAEYDAYREYTAQMSQLEGGTSGTISG 1352
 DB 1293 SMOQSTRPPLSLNSQDSISIKLTKVQAEYDAYREYTAQMSQLEGGTSGTISG 1352
 QY 121 SMOQSTRPPLSLNSQDSISIKLTKVQAEYDAYREYTAQMSQLEGGTSGTISG 180
 DB 121 SMOQSTRPPLSLNSQDSISIKLTKVQAEYDAYREYTAQMSQLEGGTSGTISG 180
 QY 1353 RSSPHSYIYIGQSSSGSISHTLEOERKGEKGLKQDGRKSPMKRGVDIVDSSGVSNN 1412
 DB 1353 RSSPHSYIYIGQSSSGSISHTLEOERKGEKGLKQDGRKSPMKRGVDIVDSSGVSNN 1412
 QY 181 RSSPHSYIYIGQSSSGSISHTLEOERKGEKGLKQDGRKSPMKRGVDIVDSSGVSNN 240
 DB 181 RSSPHSYIYIGQSSSGSISHTLEOERKGEKGLKQDGRKSPMKRGVDIVDSSGVSNN 240

```

QY 1413 EASPLDPTIEDEKSDGSKLLPGKSSERSLPTDILKLGGLRQKLPSPDESGT 1472
D 241 DASPLDPTIEDEKSDGSKLLPGKSSERSLPTDILKLGGLRQKLPSPDESGT 300
QY 1473 GNVQITPHCSKIRTKAKORCASPOEHSAPFIRTFIRAKETLSALDLKDKSSPSG 1532
D 301 EESDMP-LLKDKDKRKAQKVERVPKSPESHAERIRFIRAKETLSALDLKDKSSPSG 359
QY 1533 VSSSSSPHSHLNNAADDSOLEKNNLEDEGHSKRGMPHSISGLDPTIARMSICS 1592
D 360 VSSSSSPHSHLNNAADDSOLEKNNLEDEGHSKRGMPHSISGLDPTIARMSICS 419
QY 1593 EDKSPSECSLIASSPESWPAQKAYNLNRPSTVTNNPTANNRANONDELEGIRE 1652
D 420 EDKSPSECSLIASSPESWPAQKAYNLNRPSTVTNNPTANNRANONDELEGIRE 479
QY 1653 TSQVILRPGSPNPRAVONENIKSNAHRKRSQSSYTRLSKDSASELH-AASSESTGFEER 1711
D 480 TSQVILRPGSPNPPTTIONENIKSNAHRKRSQSSYTRLSKDPPELHAAASSESTGFEER 539

```

```

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svrtas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin C.M., Venter J.C.,
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL: AE003453; AAF46710.1; -.
DR HSSP: P42773; 11HB.
DR FLYBASE: FBgn0034615; CG10074.
DR InterPro: IPR002110; ANK.
DR InterPro: IPR000847; HTH_LysR.
DR Pfam: PF00023; ANK; 10.
DR SMART: SM00248; ANK; 10.
DR PROSITE: PS50088; ANK_REPEAT; 9.
DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE: PS00044; HTH_LysR_FAMILY; UNKNOWN_1.
KW ANK repeat; Repeat.
SQ SEQUENCE 1498 AA; 165031 MW; 7C2A888E1B64982C CRC64;

```

Query Match 26.8%; Score 2376.5; DB 5; Length 1498;
 Best Local Similarity 37.6%; Pred. No. 8.4e-146;
 Matches 552; Conservative 272; Mismatches 527; Indels 117; Gaps 36;

```

RESULT 6
QW211 PRELIMINARY: PRT: 1498 AA.
AC 09W211:
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CG10074 protein.
GN CG10074.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amaratunga P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolintsov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Charry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dooon K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J.F., Puri V., Reese M.G.,
RA Reinelt K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

```

```

QY 1 MSVLISQSVNYEEENIPALKALE-KCKVDNERNECOTPLMAEONGVEYKELK 59
D 23 MGSIGHRALLQYIDNNDISGLRAILDSRHILTDDBRENNTYLVAVAGRLAFREFLA 82
QY 60 NGANCNLEDNNNTALISAKREGHIVHELKSGSLBHRPMGCTVAMACYGRDY 119
D 83 RGDVQAEEDNNNTALISAKRNLHVVQLLDHGAEVHRDMGCTSLMAAYAGHEL 142
QY 120 VELLSHGANPSVTGLQSYVPIYAGSHADIVHLLQNGAKVNSDKYGTPLWAA 179
D 143 VRLLDKAGDGNAGH-NYHIGALLMAAGGYKDIYELLVQRAKAVGDKYGTALWAC 201
QY 180 RKHLECYVHLLMAGADVQEGANSMTALIVAKGYQSVKELIKRNPNVLTOKDGT 239
D 202 RRGVNEIVDTLKAGANVDTAGMTSWTPLVAAAGGHTCVSSILEKRNVALDKDGT 261
QY 240 ALMIASKEGHIEIVDLDAGTYVNIIPDRSGDTVLIGAVGSHVEIVRALOKYADIDR 299
D 262 ALCIASREGFODIAASLIAAGAVINIDRGADTPLIHAKAGHRTVEALLKHAIDVIO 321
QY 300 GQONKALVAVAKGNAVTRDILQCNPTETCTDGEPLIKAKRMRIEVELLDKG 359
D 322 GKDKRTAITYAVEKGTPTVYKLLATNPDESATKDGDTPLRAVRNRLEIVHLLDRK 381
QY 360 AKVASDKKGDPLVAIRGSRRLAELLRNPKDGRLLYRPNAKGETPYNDICSHOKSI 419
D 382 AKVTASDKGDPTCLHIAVARAKTIVEALLRNPKSOLLRANKKGETPYNDISLHOTTI 441
QY 420 LTDFGARHLSPETDGMVGYDYSSALADILSEPMPQPICVGLYAWQSGSKFLLK 479
D 442 LGQVFGRRLNNESEGMVGYELYSALADVISEPTTLTPTVGLYAKWGSKGFLLNK 501
QY 480 LDEKMTFAGQGTPELPFQSWLIVETLLTGGLGLVAFVPR-DINLAIAISLSTALITY 538
D 502 LRDENNNFRQNAEPPRTISGLIFTEVCLHALLITIGLSTWSAVVAVAGVLLAY 561
QY 539 IFPIVYFGREGESWMAALSTRLAHIGYLELFLKMFVNPPELPQOTTKALPYRF 598
D 562 LLLAIVRYCNGV-MDMQWAVSVQGLKRMRLRLILQVAFCHPG-PODSQAKPYRF 618
QY 599 LFTVNRKSSVGGESLAEATLTSACEREGFLATLFRVFRTE-ESQKKRKKWKTC 656
D 619 HFAEANSASPT-GDQAVVHMLAALDAIESHYGWLATRLYAFRCKLQVVGWRWRMC 677
QY 657 CLPSVFIFLFIYCGIAGITTLIAI-----RVDPKHLVNMILISAVGLAVLNORT 711
D 678 CLPIVIFELALVYVVGISLIVAYFTFADEKEKEHILV--ALYVIAVWGTLLICTHLY 735

```

QY 712 WMOVLDSLNSQRRRLHSAASKLHKLKSGFMKLCFVELMARMMAKTIDFTONOTRLV 771
 Db 736 LAVFVSLFTSHIRVLKRAV-----RSSESAPLTMGAFAVAVMTDMKCLDAFLNOOSRLV 791
 QY 772 VIIDGDACBQDKVLOMLDTRVLFs--KGPFAIFASDPHIIKAINONLSVLDNSNI 829
 Db 792 GVIDALDSCDTERLTILNNAVOGTLTSSPNRPFVLLISVDPHIAKAAENSRSLFTEGGI 851
 QY 830 NGHDYRNIVHLVFLNLSRGLSNARKFLVTS-----ATMGDITGSD--TTGDEIDTRV 882
 Db 852 GGHDFLNLVHLVFLYLNQSLRKYQAOMTALLFKRSGGDDYQTDGPTLG--HSVARKI 910
 QY 883 SONS--LGEWTKL-----GSKTALNR-----RDYRRROMOTTRTROMSPFLTKL 925
 Db 911 SNASEIISQEKRLGRARGGGKKRLSESVASTSNIHRGQNGYV-----DLRSRI 965
 QY 926 LVTEDMFSDISPOTRMRLNIVSGLRLRANOITENMDRLASWINLTQMPYRTSMIL 985
 Db 966 VLTDDYFSDVNPMSMRMLNMVITYVRLKAFQIEFSWYKLSMINLTBQMPLRASMIYL 1025
 QY 986 YLDE--TEGLPDQMTLKTMYERISKNIPTTKDVEPLELIDGDIRNEVFLSRTPLVAR 1043
 Db 1026 HHDQFNDSDNDESVSLSQSYEKRLPKATYIREAPLLEIDRERKLDALQLHKSQDLVA 1085
 QY 1044 DVKFTPLCTVNLDPKLEIITADYRAAREQINIGLAVPPLLHEGP---PRPGSYSOP 1099
 Db 1086 DLHFLPFTINLPYLKVLKEDQDTE--DEGSIVIQARPSVNTMRQEPARTTYVPS 1143
 QY 1100 ASYSSASFNPGPPGVSPHSSYSGISGPOHPYRNRAVPATGSSIL---LSSMTV 1156
 Db 1144 QAVPYQMFONEYPAN---ELRSRNLSTSTEPVPLINSPS--DSFGDDILQKTLIDLTV 1198
 QY 1157 DVVCEKRLQEGIDQNMPOCYCTIKKANINGRVLSQCNIDELKKEMAMEGWHFRSM 1216
 Db 1199 EGYISLIDRTEDM-KPALPKLAPVLRNMAINGVLAHCDMDPLKSVLGSFGHWEFLRL 1257
 QY 1217 VLEKRSVE-----SOVVPEDPRELNNSSAPV-----PHGESARRS---SHELP 1258
 Db 1258 ITTLREGERLPKQQRQOQPGALNAPSNVPMIKDVTALDAMPRESLSKKNVSIMENQ 1317
 QY 1259 LT--ELSSQTPYTLNFSFEELMTGLDEGAPRHSNLSWOSQTRRPSSIL-----N 1308
 Db 1318 VTLLEOMICGTLQLINEAYE--DVASSERPSPTGEMLAAVAQOLAPIRESEFSGSPSD 1376
 QY 1309 SODSSEISKLTDKQV---AEY-RDAFREYIAQMSQ-----EGGGSSTI---SGRSS 1355
 Db 1377 QKQYGVKISNNNNNOYLAEYNRVSSSHLSQSLSTLVGAPVGHGSGGSHLHLGNGNDL 1436
 QY 1356 PHSTYYIGQSSS-----GGSHSTLEL 1377
 Db 1437 SDSTFLDMHYDSVFGGGGGGYHRASQ 1464

RESULT 7
 Q20109 - PRELIMINARY; PRT; 1398 AA.
 AC Q20109:
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE F36H1.2 protein.
 GN F36H1.2.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Chromadorea; Rhabditida; Rhabditoidea;
 CC Rhabditidae; Peleoderinae; Caenorhabditis.
 CX NCBI_TaxID=6239;
 RN [1]
 RN SEQUENCE FROM N.A.:
 RA Kershaw J.K.:
 RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
 RP [2]
 RP SEQUENCE FROM N.A.

RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode C.elegans: A platform for
 RT investigating biology."
 RL Science 282:2012-2018(1998).
 DR EMBL; 268760; CAA92996.1; -.
 DR HSSP; 000420; IAMC.
 DR InterPro; IPR002110; ANK.
 DR InterPro; IPR001412; tRNA-synt_1.
 DR Pfam; PF00023; ank, 10.
 DR PRINTS; PRO1415; ANKTRIN.
 DR SMART; SM00248; ANK; 9.
 DR PROSITE; PS00178; AA_tRNA_LIGASE_I; UNKNOWN_1.
 DR PROSITE; PS00088; ANK_REPEAT; 7.
 DR PROSITE; PS00297; ANK_REPEAT_REGION; 1.
 KW ANK repeat; Repeat.
 SQ SEQUENCE 1398 AA; 156568 MW; D5080A9E89AB1BEC CRC64;

Query Match 20.9%; Score 1848.5; DB 5; Length 1398;
 Best local similarity 31.5%; Pred. No. 2,3e-111;
 Matches 461; Conservative 272; Mismatches 459; Indels 271; Gaps 37;

QY 34 RNECGTPLMLAEOGNEIYVELLKNANCNLELDN--WTALISAKREGHIVYELL 91
 Db 58 RNANESLTLTVAVRSGNAVAQAQLDLPDA-IDETDNEGSALLNMAHGHVIVRLLI 116
 QY 92 KSGASLEHDMGKWTALMACYKGRDVEYELLSHGAPSVYGLQYSYPIIWAARGHA 151
 Db 117 DMGASVDPQDLNKGSPMLMAVYKKNLDVVDLVN----- 150
 QY 152 DIVHLLONGAVNCSKYGTTPLVWAARKGHLDECVKHLAMGADVDEGANSMTALIVA 211
 Db 151 -----AKVNMCDKFGSTALIMAAKRGHLPVQQLINSGAEVDAVMYSTALMLA 200
 QY 212 VNGGTYQSVELLKRPNVNLTDKGNATMALASKEGHEIYODLLDAGTYVINPDRCSD 271
 Db 201 TRGNFIOVVELLTTRPNVNVADONGTLAAGAADGYADICESLINSGAVNOCGRGN 260
 QY 272 TVLIGAVRGHEIYRALLQKADIDIRGDNKTLALYNAVEKGNATVTRDILQCPDEI 331
 Db 261 WILTSVRSKGNALYRMILDKRADINDQDSERKTLHLAIDKSPNDIAYILLEKPNLEL 320
 QY 332 CYRQDEPLIKATKKNIEVEVELLDGAKVSAVDKGDPLHVAIRGRSRLAELLRN 391
 Db 321 KKKDGETPLLRKAKCHVHLCTYMSFGKLAAYVNCGMNLHLALRRASRLQALLSN 380
 QY 392 PKDGLLYPNKAGETPYIDCSHOKSILTOFGARHLSPTDGD---LGYDIYSAL 448
 Db 381 PSDSRLLYRPNKLGOTPYSIDLSNPQPLPLIFG-----PIDAEKMDTAMGYDYSVL 435
 QY 449 ADLSEPTMQPPICVGLYQMGSGSKFLKLEDEKTFPAQQTPL-FOFSMLIVFTL 507
 Db 436 ADIVCEPSLSPLTIGLYAKMGSKSALLAKLEMAHSSRMWDGVSLSYFALFPIAF 495
 QY 508 LILGGILVAFAPV---DTNLAIASLSFLALIFYF-----IVYFGRRGESNNMMA 560
 Db 496 LFGGMSLFTMLAISNSVTAYILMSVFLIFLIFICSILVAVVYGYDRK---NMVYSMD 552
 QY 561 LSTRLARHIGYLELFLKMFVNPPELPEDTALVPRPLFTNYNLSVSGGTSIAEMA 620
 Db 553 IANFEARVSRRLRVNLTLLHAP-MNSDSASMPVSFLFADYHRLSRISGGQALAKIVA 611
 QY 621 TUSDACEREGFLATLRFVFRTEESQKKKKCTCCLPSFY-----IFLIVGCTIAGI 675
 Db 612 TLFEAEHTFQVLPVRLFCCKMPYPYGGISGLRRRGQVHVILLVAVFLIMAGVPGV 671
 QY 676 TLIAIFRVDPKHLTVNALLISIASVGLAFVNLNCRTWQVLDLSNQRRLHSAASKLH 735
 Db 672 WILTS--DRDPNNFN--LFTAIATFCGFMAIATYPLALIMYSWNVRRRRVNAARAH 726
 QY 736 KLSGEGFMKLVCEVELMARMMAKTIDFTONOTRLVVIIDGDACBQDKVLOMLDTRV 795
 Db 727 KLFEGLMKQLQTEVDLADMRSLDAPFRSHTRLVVAVVGDINCEQERAVQVOTLALIELL 786

DT 01-JUN-2002 (Tremblrel. 21, last annotation update)
 DE Hypothetical protein Al12748.
 GN Al12748.
 OS Anabaena sp. (strain PCC 7120).
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
 NC NCBI_TaxID=103690;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21595285; PubMed=11759840;
 RA Kaneko T., Nakamura Y., Molk C.P., Kuritz T., Sasamoto S.,
 RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
 RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
 RA Yasuda M., Tabata S.;
 RT "Complete genomic sequence of the filamentous nitrogen-fixing
 RT cyanobacterium Anabaena sp. strain PCC 7120.";
 RL DNA Res. 8:205-213(2001).
 DR EMBL: AP003590; BAB74447.1; -.
 DR InterPro: IPR002110; ANK.
 DR Pfam: PF00023; ank. 13.
 DR PRINTS: PR01415; ANKYRIN.
 DR SMART: SM00248; ANK. 13.
 DR PROSITE: PS50088; ANK_REPEAT. 11.
 DR PROSITE: PS50297; ANK_REPEAT_REGION. 1.
 DR Hypothetical protein; Complete proteome.
 KW SEQUENCE 426 AA; 43800 MW; CF0895D43CAA03 CRC64;
 SQ
 Query Match 6.3%; Score 555; DB 16; Length 426;
 Best Local Similarity 37.2%; Pred. No. 8.6e-28;
 Matches 149; Conservative 67; Mismatches 177; Indels 8; Gaps 6;
 QY 9 VINVEEENIPALKLEKCKDYDERNECGQTPMLAEGQNEIYKELKNGANCNL-E 67
 DB 8 LKVAKSGDKLGLALLAAGVVDICDRDSTTALMPANIGYEIYASLLDGGANVNLAR 67
 QY 68 DLDNMTALISASKEGHIHVEELKSGASLEHDMGWTALMAWYKGRDVELLSHG 127
 DB 68 KRYGLTALMLAASQVDIYQLLISGAANVANNEDGSTALMAALKGNEVARVLLAAG 127
 QY 128 ANPSVYGLQSYVPIIIMAAARGHADIYHLLONGAKYNCSDKGTPIYVAAKRGHECY 187
 DB 128 ADVNITDKDDT-ALKLAVERGGAAYVQLLIPSGADNCEDEGETLLMLAADSQGDVY 186
 QY 188 KHLAMGADYDEGANSMTALIVAVKGYTSQYKEILKRPNNVLTDKDNTALMASKE 247
 DB 187 QVLLAAGVDYNEQDQSTGLLAAYAGNAIKIILDRADVNHODQDESALHATVE 246
 QY 248 GHEIYVDLLDAGTYVNIIPRSGDTVLIGAVRGHVEIVALLQKYADIDIRGOD-NKTA 306
 DB 247 GYVDYVGVLLINGQANTQIKKRLGTPPLVALQGHDIYETLL-KYG-ANVHGDNIGETP 304
 QY 307 LYMAVEKGNATMWDILQCNPDIEICKDEETPLIKATKRNIEYVELLDKAKYSAVD 366
 DB 305 LTLLAASQGHYATYVILLDYCANANIPASDKRTALIKTEENHGVQLLLAKGANVYOD 364
 QY 367 KKGDTPLHVAIRGSRRLAELLRNPKDGRLLYRPNKAGET 407
 DB 365 SVGATALIMAAASGGYNNVQILLLEGADTNL---KNNGGT 402
 RESULT 11
 Q99407 PRELIMINARY; PRT; 1856 AA.
 AC 099407;
 DT 01-MAY-1997 (Tremblrel. 03, Created)
 DT 01-MAY-1997 (Tremblrel. 03, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Ankyrin.
 GN ANK1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 NC NCBI_TaxID=9606;

RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97382244; PubMed=9235914;
 RA Gallagher P.G., Tse W.T., Scarpa A.L., Lux S.E., Forget B.G.;
 RT "Structure and organization of the human ankyrin-1 gene. Basis for
 RT complexity of pre-mRNA processing.";
 RL J. Biol. Chem. 272:19220-19228(1997).
 DR EMBL: U50133; AAB47805.1; -.
 DR EMBL: U50093; AAB47805.1; JOINED.
 DR EMBL: U50092; AAB47805.1; JOINED.
 DR EMBL: U50094; AAB47805.1; JOINED.
 DR EMBL: U50095; AAB47805.1; JOINED.
 DR EMBL: U50096; AAB47805.1; JOINED.
 DR EMBL: U50097; AAB47805.1; JOINED.
 DR EMBL: U50098; AAB47805.1; JOINED.
 DR EMBL: U50099; AAB47805.1; JOINED.
 DR EMBL: U50100; AAB47805.1; JOINED.
 DR EMBL: U50101; AAB47805.1; JOINED.
 DR EMBL: U50102; AAB47805.1; JOINED.
 DR EMBL: U50103; AAB47805.1; JOINED.
 DR EMBL: U50104; AAB47805.1; JOINED.
 DR EMBL: U50105; AAB47805.1; JOINED.
 DR EMBL: U50106; AAB47805.1; JOINED.
 DR EMBL: U50107; AAB47805.1; JOINED.
 DR EMBL: U50108; AAB47805.1; JOINED.
 DR EMBL: U50109; AAB47805.1; JOINED.
 DR EMBL: U50110; AAB47805.1; JOINED.
 DR EMBL: U50111; AAB47805.1; JOINED.
 DR EMBL: U50112; AAB47805.1; JOINED.
 DR EMBL: U50113; AAB47805.1; JOINED.
 DR EMBL: U50114; AAB47805.1; JOINED.
 DR EMBL: U50115; AAB47805.1; JOINED.
 DR EMBL: U50116; AAB47805.1; JOINED.
 DR EMBL: U50117; AAB47805.1; JOINED.
 DR EMBL: U50118; AAB47805.1; JOINED.
 DR EMBL: U50119; AAB47805.1; JOINED.
 DR EMBL: U50120; AAB47805.1; JOINED.
 DR EMBL: U50121; AAB47805.1; JOINED.
 DR EMBL: U50122; AAB47805.1; JOINED.
 DR EMBL: U50123; AAB47805.1; JOINED.
 DR EMBL: U50124; AAB47805.1; JOINED.
 DR EMBL: U50125; AAB47805.1; JOINED.
 DR EMBL: U50126; AAB47805.1; JOINED.
 DR EMBL: U50127; AAB47805.1; JOINED.
 DR EMBL: U50128; AAB47805.1; JOINED.
 DR EMBL: U50129; AAB47805.1; JOINED.
 DR EMBL: U50130; AAB47805.1; JOINED.
 DR EMBL: U50131; AAB47805.1; JOINED.
 DR EMBL: U50132; AAB47805.1; JOINED.
 DR HSSP: Q00420; IAWC.
 DR InterPro: IPR002110; ANK.
 DR InterPro: IPR000488; Death.
 DR InterPro: IPR000906; ZUS.
 DR Pfam: PF00023; ank. 23.
 DR Pfam: PF00531; ZUS; 1.
 DR Pfam: PF00791; ZUS; 1.
 DR PRINTS: PR01415; ANKYRIN.
 DR SMART: SM00248; ANK. 21.
 DR SMART: SM00005; DEATH. 1.
 DR SMART: SM00218; ZUS; 1.
 DR PROSITE: PS50088; ANK_REPEAT. 20.
 DR PROSITE: PS50297; ANK_REPEAT_REGION. 1.
 DR PROSITE: PS50017; DEATH_DOMAIN; 1.
 KW ANK repeat; Repeat.
 SQ SEQUENCE 1856 AA; 203466 MW; 4A4E3A794DD6401F CRC64;
 Query Match 6.3%; Score 555; DB 4; Length 1856;
 Best Local Similarity 20.6%; Pred. No. 1.1e-26;
 Matches 395; Conservative 269; Mismatches 644; Indels 612; Gaps 74;
 QY 21 LKALLEKCKDYDERNECGQTPMLAEGQNEIYKELKNGANCNLELDNMTALISASK 80
 DB 61 VRELIVYGANVNAQSGKFTPLYMAAQENHLEVFLENGANQNVATEDGETPLAVALQ 120

```

OY 81 BSH----- 83
DB 121 QHEHVAHNLNYGKGVRLPALHIAARNDRTAVALLONDPNPDLSTGFTPLHIA 180
OY 84 -----IHVEELLSGASLEHRDMGWTALMACYKGRDVELLSHGANSPTGLDYS 138
DB 181 AHENLVANAOILLNNGASVNFPTPONGITPLHIASRGVIMVRLLDGACIE-TKTKDE 239
OY 139 YPIITWAGRGHADIIVHLLONGAKVNSDKYGTTPVWAARKGHECVKHLAMGADV 198
DB 240 LTPHCAARNGHVRISEILDHGAPICAKTNGSLPHMAAQGHLCVRLLOYDAEID 299
OY 199 -----OECA-----NSMTALIVAGGYTQVKEILK 225
DB 300 DITDLHPLVHAHCHHRAKAYLLDKGAPNSBALGFTPLHIAKQVAVARSLLOYG 359
OY 226 RNPVNLTDKDGNTALMIASKEGHEIYODLLDAGTVNIPDRSGDVLIGAVRGHVEI 285
DB 360 TGAIDAVTESGLFPLHVASFMGHLPIYKNLLDGGASPNVSNVETPLHMAARAGHVEV 419
OY 286 VRALLQKADDIRGODNKATLYANAVEGNATWVRDILQCPDTEICTKOGETPLIATK 345
DB 420 AKYLLONKAKANAKADOTPLHCAARIGHNMVKTLENNANENLATAGHTPLHIAAR 479
OY 346 MNIEVVELLDKGAKEYSAVDKGDTPHVAIRGRSRLAELLRNPKDGLLYRPKAG 405
DB 480 EGHVETVALLEKESQACMTKGTPLHVAAGKAVVALLLE-RDA-----HPNAG 533
OY 406 E---TPYIDCSHO-----KSIL-----TOIFGAR----- 427
DB 534 KNGLPPLHVAHNNLIDIVKLLPRGSGSPHANNGYTPLHIAKQVAVARSLLOYG 593
OY 428 -----HLSPETDMDLGYDLYSSALADLISEPTMP-----PIC-- 462
DB 554 SANASVOCVTPPLHIAOEGHAEMVALLSKOANGNNGKSGPLPLHVAOEGHVPADV 653
OY 463 -----VGLYAOWSGK--SPLIKLED-EMKTFAGOOQTEPLPFSWL 501
DB 654 LIKHGVMDATTRMGYTPPLHVAHSHYGNIKVFLLOQADVNAKTIGY--SPLHQAQ 711
OY 502 --IYFLLILCGGLVFAFPVD-----TMLAIALISFLALYIFFIYFGGRGE 553
DB 712 GHTDVLVLLKNG-----ASPNEVSDGOTPLLAIAKRLGYISVDVLYV----- 756
OY 554 SMNAMALSTRLARHIGYELLFKLMFYNPEPEQTTPKALPVFLTDVNRLS-SYGE 612
DB 757 TDESFVLYSDKHR--MSPETVDELIDVSEDEGTAHITIMGEELISKRAEROSRDVEE 815
OY 613 TSLAEMIATLSDACEREGFLATRLFR-----VETEE--SQGKKMKKTCCLPSFV 662
DB 816 KELIDFVPLDOVE-----SPAIPRIPCAMPETVYIRSEEGQAKVEDEBLSIPS-- 867
OY 663 IFLFVGTIINGITLLAIFRVDPKHLVNALISIASVGLAFLNCRIMQVYDLSLNS 722
DB 868 -----SPATEISDNI--SPVASPVHGFVLSF-----MYDAGS 900
OY 723 QKRRLHSAASKLHKLSGFMKVLKCEVELARAKTIDSTOQUTRLVYIIGDLACEQ 782
DB 901 MRGRHN-----GLRVIIP-----PRICAPRTITCLV----- 929
OY 783 DKVLQMLDTVAVLESKGFPIAFASDPHIIKAINOMINSLVLRDINSINGHDYMN-IVHL 841
DB 930 --KPKLSTPPL-----AEEGLASRIIALGPTGAQFSPVIVEI 968
OY 842 PVPL-NSRGLSNAKFLVTSATNDITCSDTGTQOEOTDRVYSONSIGEMTKIGASKTALN 900
DB 969 PHFASHGSG--DRELVVLRSENGSV-----WKHRSRGYSYLDQILN--GMDELG 1016
OY 901 RRDYRRRORORTITROMSFDLTKLVY-----EDWFSISPOTMRLNITVSGRLLRAN 957
DB 1017 SLELEKRRKCRIT--TDPLLYIVMSRLCQD-YDITGPEG-----GSLSKVLPLY 1066

```

```

OY 958 QITENMDRLASWINLTEQMPYRTSWLILVLEETEGLPDQMTK-----TMYER 1005
DB 1067 QATFENAVATKRLALQ-----AQVPDELVTKLGNQATFSPVIVEPR 1112
OY 1006 ISK-----NIPTKVEPLEIDGDIRNEFEV-----LSSRTPVVA 1042
DB 1113 RRFHNRPIGIRLPLPSWMDNPNRDSGEGTTLRLCSYIGGDAQWEDITGTXIYVA 1172
OY 1043 RDKVTLPLCTVINIDPKRELIADYRARBOINIGLAPPLPHHGPPRPGYOPASV 1102
DB 1173 NECANF--TTNV--SARFMLSOPRTAAVAVATLTKEL-----TAVPYMAF 1217
OY 1103 CSSASFNGPPGVSPQPHSSVYSGLSQPHFYRAAVPATGSSLLSMTVYVCEK 1162
DB 1218 VIFAKMNDREG-----RLCYCMTDDKVDYT 1244
OY 1163 LRO-----IEG-----LDONMPOYCTTIRKAN-----INGRYL 1191
DB 1245 LEQHEHVEVARSROIEVLEGMSLFAELSGNLV-----YKKAQOGRSFHFGSPENRLA 1299
OY 1192 SOCNIDELKE-----MANNED--WHLRSMVLEKRSVESOVVPEPR----- 1233
DB 1300 MPYKVRDSSREPGSLFLRKAKMYEDTOHILCHLITMPCKAGSGAEDRRRTPLAL 1359
OY 1234 --FLNENSGAPVPHEGSAAR-----SHTLPLTFLSQTPTLNFSPLENTIGLD-- 1283
DB 1360 RYSLISESTGSLSGTEQAEKMAVISEHLIGSMELARE-----LOFSYEDINRLVENP 1415
OY 1284 ----EAPRHSNLSWOSQTRPPSSLNSODSSIFSLTKDVQA-----EYRD 1329
DB 1416 NSLLEQSVALLNL--WYIREGONAMENLYTALOSIDREIVNMLDESGROSRLPKDRH 1474
OY 1330 AVEEYIAQMSOLEGGGSGSTISGRSSPHSTYIIGSSGSGSHSTEDRGREGELKQED 1389
DB 1475 TDRDYSLSPSQMNG--YSSLODELSP-----ASLCAALSPRLAQ----- 1514
OY 1390 GRKSFMRKGDVYDSSGVSSTNEASPLDPI-----TEED--EKSPDS--GSKLPGRKS 1440
DB 1515 -----YNEVAVDAITPLAETHDTMLEMDSMQVMSGLRPLVLT 1554
OY 1441 SERPSLFQDILKKGGLRQKPLSPDEDSGGRVOITPHCSKMTITRKLAKQRECAP 1500
DB 1555 AEDSSL-----BCSKAEDSDATGH-----EMKLEGLAISEPRGP 1588
OY 1501 QEHSADPIRPFIRAKETLSDA---LIDKDDSSDGSVRSNESSPNHLEADDSOLEKA 1557
DB 1589 ELGSL-----LYEDDTVSDATNGLDLLE--QEBGORSSEKLPGRKODDAGACQ--DSE 1642
OY 1558 NLELEDEHSGKRGKMPHSLSGLOPPIIARMSIGSEBDK-----KSPSECSLIASSPERSW 1612
DB 1643 NEVSLVSGHQROGARITHS-----PTVQVTERSQDRLOQMDADGSIYSYLODAAGSWM 1696

```

RESULT 12

061307 PRELIMINARY; PRT; 1943 AA.

AC 061307; 061305; 061306; 061308; 061309; 061310; 008866; 008867;

DT 01-JUN-1998 (TREMblrel. 06, Created)

DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)

DE 01-DEC-2001 (TREMblrel. 19, Last annotation update)

DN Ankyrin 3 (Ankyrin 3) (epithelial ankyrin) (Ankyrin-3).

GN ANK3.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI; Taxid=10090;

RN [1]

RP SEQUENCE FROM N.A. (ISOFORMS 1-6).

RX STRAIN=C57BL/6J; TISSUE=KIDNEY;

RC MEDLINE=95340633; PubMed=7615634;

RA Peters L.L., John K.M., Lu F.M., Elchler E.M., Higgins A., Yialamas M.,

RA Turtzo L.C., Otsuka A.J., Lux S.E.;

RT "Ank3 (epithelial ankyrin), a widely distributed new member of the

RT ankryn gene family and the major ankryn in kidney, is expressed in
 RT alternatively spliced forms, including forms that lack the repeat
 RT domain." J. Cell Biol. 130:313-330(1995).
 RN [2]
 RN SEQUENCE FROM N.A. (ISOFORMS 7 AND 8).
 RC STRAIN-C3H/HEJ; TISSUE-BONE MARROW;
 RX MEDLINE-97213781; Pubmed-9060470;
 RA Hoock T.C., Peters L.L., Lux S.E.;
 RT "isoforms of ankryn-3 that lack the NH2-terminal repeats associate
 RT with mouse macrophage lysosomes.";
 RL J. Cell Biol. 136:1059-1070(1997).
 CC -1- FUNCTION: ANKRYNS ATTACH INTEGRAL MEMBRANE PROTEINS TO CYTO-
 CC SKELETAL ELEMENTS; THEY BIND TO THE ERYTHROCYTE MEMBRANE PROTEIN
 CC BAND 4.2, TO NA-K ATPASE, TO THE LYMPHOCYTE MEMBRANE PROTEIN GP85,
 CC AND TO THE CYTOSKELETAL PROTEINS FODRIN, TUBULIN, VIMENTIN AND
 CC DESMIN. ERYTHROCYTE ANKRYNS ALSO LINK SPECTRIN (BETA CHAIN) TO
 CC THE CYTOSOLIC DOMAIN OF THE ERYTHROCYTE ANION EXCHANGE PROTEIN;
 CC THEY RETAIN MOST OR ALL OF THESE BINDING FUNCTIONS.
 CC -1- SUBCELLULAR LOCATION: PLASMA-MEMBRANE ASSOCIATED. ISOFORMS 7 AND 8
 CC ARE ASSOCIATED WITH THE LYSOSOMAL MEMBRANE.
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 8 ISOFORMS, 1 (SHOWN HERE), 2, 3,
 CC 4, 5, 6, 7 AND 8; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: MAJOR FORM OF ANKRYN IN KIDNEY AND OTHER
 CC EPITHELIAL TISSUES. ALSO EXPRESSED IN MACROPHAGES, MEGAKARYOCYTES,
 CC LEYDIG CELLS, CARDIAC, SMOOTH AND SKELETAL MUSCLE, INITIAL
 CC SEGMENTS OF AXONS, AND NODES OF RANVIER.
 CC -1- SIMILARITY: CONTAINS 24 ANK REPEATS.
 DR EMBL; LA0631; AAB01603.1; -;
 DR EMBL; LA0632; AAB01605.1; -;
 DR EMBL; U89275; AAB58381.1; -;
 DR EMBL; LA0631; AAB01602.1; -;
 DR EMBL; LA0632; AAB01604.1; -;
 DR EMBL; LA0632; AAB01606.1; -;
 DR EMBL; LA0632; AAB01607.1; -;
 DR EMBL; U89274; AAB58380.1; -;
 DR HSSP; P55273; IBI8.
 DR MGD; MGI:88026; Ank3.
 DR InterPro: IPR002110; ANK.
 DR InterPro: IPR000488; Death.
 DR InterPro: IPR000906; ZU5.
 DR Pfam; PF00023; ank; 24.
 DR Pfam; PF00531; death; 1.
 DR Pfam; PF00791; ZU5; 1.
 DR PRINTS; PR01415; ANKRYN.
 DR SMART; SM00248; ANK; 21.
 DR SMART; SM00005; DEATH; 1.
 DR SMART; SM00218; ZU5; 1.
 DR PROSITE; PS50088; ANK_REPEAT; 21.
 DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
 DR PROSITE; PS50017; DEATH_DOMAIN; 1.
 KW Cytoskeleton; Alternative splicing; Repeat; ANK repeat;
 KW Multigene family.
 FT DOMAIN 1 826
 FT DOMAIN 23 810
 FT DOMAIN 856 1437
 FT DOMAIN 1438 1943
 FT REPEAT 23 55
 FT REPEAT 56 88
 FT REPEAT 89 121
 FT REPEAT 122 154
 FT REPEAT 155 183
 FT REPEAT 184 216
 FT REPEAT 217 249
 FT REPEAT 250 282
 FT REPEAT 283 315
 FT REPEAT 316 348
 FT REPEAT 349 381
 FT REPEAT 382 414
 FT REPEAT 415 447
 FT REPEAT 448 480

FT REPEAT 481 513 ANK MOTIF 15.
 FT REPEAT 514 546 ANK MOTIF 16.
 FT REPEAT 547 579 ANK MOTIF 17.
 FT REPEAT 580 612 ANK MOTIF 18.
 FT REPEAT 613 645 ANK MOTIF 19.
 FT REPEAT 646 678 ANK MOTIF 20.
 FT REPEAT 679 711 ANK MOTIF 21.
 FT REPEAT 712 744 ANK MOTIF 22.
 FT REPEAT 745 777 ANK MOTIF 23.
 FT REPEAT 778 810 ANK MOTIF 24.
 FT VARSPPLIC 1 805 MISSING (IN ISOFORM 7 AND ISOFORM 8).
 FT VARSPPLIC 1 849 MISSING (IN ISOFORM 2 AND ISOFORM 3).
 FT VARSPPLIC 833 853 MISSING (IN ISOFORM 7 AND ISOFORM 8).
 FT VARSPPLIC 834 855 VRKASPELSDGEYSIGEEG -> GDCIKFWIKPKVOEV
 FT VARSPPLIC 850 855 LVK (IN ISOFORM 5).
 FT VARSPPLIC 850 855 SDGEG -> MALPRS (IN ISOFORM 2 AND
 FT VARSPPLIC 852 852 ISOFORM 3).
 FT VARSPPLIC 852 852 G -> GDCIKFWIKPKVOEVLYKS (IN ISOFORM 4
 FT VARSPPLIC 1569 1764 AND ISOFORM 6).
 FT VARSPPLIC 1569 1764 MISSING (IN ISOFORM 3, ISOFORM 4 AND
 FT VARSPPLIC 1569 1764 ISOFORM 7).
 SQ SEQUENCE 1943 AA; 212013 MW; B23C4B6F82904804 CRC64;
 Query Match 6.2%; Score 551; DB 11; Length 1943;
 Best Local Similarity 20.2%; Pred. No. 2.2e-26;
 Matches 426; Conservative 289; Mismatches 666; Indels 724; Gaps 85;
 14 EEENIPALAKLEKCKDVERNECCOTPLMLAEGNVEIVKELKNGANCLDEDLDMWT 73
 66 KEGHEVVSSELLDREANVDAATKKGNTALHISLAGEVAVLVNAGNNAOSQNGFT 125
 74 ALISAKKEGHIHIVEELKSGASLEHRDNGWTALMACYKGRDVELL----- 124
 126 PLYMAQENHLEVRRLDNGASQSLATEEDGTPPLVALQGHQDVSLLENDTKGYR 185
 125 -----SGANPSYGL----- 135
 186 LPALHIAARKDDTKAALLLDNDYADVESKSGFTPLHIAHYGINVATLLNRAAVD 245
 136 ---QVSYPPLIIMAGGHADIYHLLONGAKVNCSDKYGTPLVWAARK----- 182
 246 FTRANDITPLHVAASKKGNAMVKLLIDRAKIDAKTRGSLTPLHGARGHGEVEMLLD 305
 183 -----HECVKHLAMGADVDOEGANSMTALIAVKGGYQS 219
 306 RSAPILSKTKNGLSPHMAVTOGDHLCVOLLQHNVPVDYDNDYLLALHVAHCGHYKV 365
 220 VKETLRNPNVNLTDGNTALMASKE----- 247
 366 AKVLDDKKASPNKAKALNGFTPLHIAKKNRIRVMEILLKNGASIQAVTESGLTPLHVAE 425
 248 -GHIEIYODLDGATVNIIPDRSGDVLIGAVGCHVEIVRALQKYADIDIRGDNKTA 306
 426 MGVNIVISQMLHMGASPNNTNNGETALIMARSQAEVYRVLVDGQAVEKAKADDP 485
 307 LYMAVEKGNATVNRDILQCNPDTEICTKDETPPLIKATMRNIEVELLDGAKYSAVD 366
 486 LHISARIGADIYQOLQOGASPNNAATSGYPLHIAAREGHEDVAAPFLDNGASISITT 545
 367 KKQDTPLHVAIRKSRRLAELLR-----NPKDGRLL----- 398
 546 KKGFTPLHVAAKYKGLVAVSLLLQKSASPDAGSGSLTPLHVAHYADNCKVALLLDGGA 605
 399 --YRPKAGETPYNIDC-SHOKSILTOI--FGAR-----HLSPTEDGDMILGY 441
 606 SPHAARKNGTTPHIAAKKNQMDIATSLDEYGADANAATVROGIAASHLAAGSHVMSL 665
 442 DLYSSALADILSEPTMOP-----PICVGLIYAW 469
 666 ILSRNVNVLNKSGLTPLHIAQEDRVNVAEVLVNOGAVDAQTGMGYTPLHVGCH--Y 723
 470 GSKR--SFLK---KLEDEKTFAGQOTETPLQFSLLIYELTL-----LLCGSLGLV 516

```

Db 724 GNKIYVFLQSHSAKVNATKNGYTALHQAQGHTHIINVLQNNASPNELTVNG----- 779
QY 517 FAFPDVTLNLAISIFLALYIFVI----- 544
Db 780 -----NTALAIARRRGYISVDTLKVVTEIWTITTEKHMNPETMNEVLDMDEY 834
QY 545 -----YFGREGESMMAMALSTRLARHIGYLELFKLMFVNPPELPEOT 591
Db 835 KASAPKLSDEEYISDGEED-----AIGEDTKYLIG-----PODLKEGD 877
QY 592 KALPVAFLETTYDNRSLSSVGETSLAEMIALTSDACEREFGLATRLFVFPREESQKK 651
Db 878 DSLPPE-----GYG--FSLGARASLR-----SFSDRSTYLNSSVAK----- 915
QY 652 MKKTCCLDSFVIFLIVGCIAGTLLAIFRVP-KHLT-----VNAILISIASVYG 702
Db 916 -----DSMIEELVPSKEQHLFTREDFSDSLRHSWMAADTLDNVNLVSSPVHSFL 968
QY 703 LAFVLCRFRWQVLDLSLNSQ-----RK-----RLMSASKLHL-----KSE 740
Db 969 VSEVDARQ-----GSMGSRHGMRIIPPKCTAPTRICRLVKRKLANPPPMVEGE 1023
QY 741 G-----FMKVLKCEVELMARAKTIDSTFONO-----TRL 770
Db 1024 GLASRLVEMGPAQFLGPVIVEIPFGSMGKRELIVLASENGETWKEHQPSKMDL 1083
QY 771 VTIIDGACEDQKYLQMLDVRV--LFSKGFPIAFASDPHIIKAINQNLNSVLRDSN 828
Db 1084 AELNGMD--BELDSPEELGTKRICRITK-DFQYFA-----VYSRIKQESNQIGPBG 1135
QY 829 INGHYMNRIYALPFLNSRGLSNARKFLV-TSANTNGITCSDTTGTQ-----EDTDREV 882
Db 1136 I-----LSTTYPLVASPEGLTRIRIVGLOAQVPEETVAKI 1175
QY 883 SONSIGEMTKLSKTALNRDPT--RRROMRTITROMSFDFLKL--VTEDMFSDIP 937
Db 1176 -----LGNKATFSPIYVEPRRRKFKPIT-MTIVPPSGSGVSNKGKDATP 1223
QY 938 QTRMLNIVSYTG-----RLRANOITFNMNDRLASWINTLEQMYRTSWL-----I 984
Db 1224 NL--RL--CSITGTPAWMEDITGTPPLFKIDCVSFTTNVARSF-----WLADCHQV 1274
QY 985 LYLEETEGLPDQMTLKTW-YERISKNIPTTKDVEPLEIDODINF----- 1029
Db 1275 L--ETVGLASQILRELICVYMAKFVFAFNTNP--VESLSLFCFMTDTRVDKLEQ 1328
QY 1030 -----EVLSSRTPVLVARDVKTPLCTVNLDPKLR--ELIADYARARE-----QINIG 1077
Db 1329 ENFEVARSKDIEVLEKPI--YVDCYGNLAPFLKGGQQLVFNFSKRENLPRESIKIRD 1386
QY 1078 LAYPLP--LHSGPPRPPSGYQSPASVCSASFNGPFGGVVSPQPHSSYVSGLSGPQ 1133
Db 1387 TSQPCGRLSFLKE--PATTGKLPOTA-VCN-----LNTILPA 1421
QY 1134 HPFYRAAVPATGSSLL-----SSMTVDVVCER--LRQIEGLDQNMMPQYCTTIK 1182
Db 1422 HKKAEKADROSFPASLARKRYSTLTPBSMSPOSCEBTDIRMAIVAD-HLGSLWELAR 1480
QY 1183 KANINGRVLSCNIDELKEMANF--GDW-----HLFSWVLEMRVESQVVP 1229
Db 1481 ELNFSVDEINOIREVENPNSLISQSFMLKKWVRDGNKATTDALTSVLITINRDIYTL 1540
QY 1230 EDP-----RFLNENS--SAPVPHGESARRSHHTELPLTELSSQTPYTLNFSFEE 1276
Db 1541 EGPFLFDYGNISGTRSFADENNVPDPV-----DGH--PSFOVELFPMGLIYMT--P 1587
QY 1277 LNTIGLDEGAPRHNSLMSQSTRTPSLSS-----LNSDSSSTEIS 1317
Db 1588 PNPPOQDD--HPSDISISPFRTPSRLSDGLVPSOGNIEHPTGPPVYTAEDTSLSDS 1644
QY 1318 KLVDKV--QAEYRDAYREVIQMSQLEGGTSGSTISG-----RSSP--HSTYIYGOS 1365
Db 1645 KMDOSVIVTDPADPLDVDESQKLDLQSECAQCMASVPGILPNDGRQAEPLRPQTRKVGMS 1704

```

```

QY 1366 SSGGSHSTLEQERKKEGELKQEDGRKSFIMKRGVITYSSGVS7NEASPLDPITEDE 1425
Db 1705 SE-----QOERKSG-----PDEVEYEDKV 1724
QY 1426 KSDQSSKLLPGKSSRPFLFOTDLKKGGLRYOKLPSDEDESGTGRVQITPHCSKI 1485
Db 1725 K-----SLFE-DIQE--EVEAEEMEDQOAMINRYO-----RAEL 1758
QY 1486 RTRKRAKORECASPQESAPERTFIKAKEYSDLLDKKSSDSGVSBN----- 1536
Db 1759 AMSLSAQWQETPSGSLSPQAQR-----LTGGLDLRLDSSDQANDSTSYLTGER 1811
QY 1537 ---ESSPNHSIHNEADDQSLERANTIELEDESGSKRGK-----PHSLGIDPPIA 1586
Db 1812 GKIEANGNHT-----AEVIEAKAKYPFEPSONDICKOSIKENLAKRTGCGTEEPV-- 1864
QY 1587 RMCISEDKSPSECS--LIASSPESWPACQAYMLNTPSTVTLNNTAPTRANQNF 1644
Db 1865 --SPLTAYQKSLBETSKLYIEDAPKPCVPYMKKM-----TRTADKARLNL 1910
QY 1645 DEIEG 1649
Db 1911 QEEBG 1915

```

```

RESULT 13
Q13768
ID 013768 PRELIMINARY; PRT: 1719 AA.
AC 013768;
DT 01-NOV-1996 (TEMBUREL, 01, Created)
DT 01-NOV-1996 (TEMBUREL, 01, last sequence update)
DE 01-JUN-2002 (TEMBUREL, 21, last annotation update)
DE Alt. ankyrin (variant 2.2).
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RC SEQUENCE FROM N.A.
RC TISSUE=HEMATOPOIETIC;
RX MEDLINE=90158830; PubMed=2137557;
RA Lux S.E., John K.M., Bennett V.;
RT "Analysis of cDNA for human erythrocyte ankyrin indicates a repeated
RT structure with homology to tissue-differentiation and cell-cycle
RT control proteins."
RL Nature 344:36-42(1990).
DR EMBL: X16609; CAA34611.1; -
DR HSSP: Q00420; IAWC.
DR InterPro: IPR002110; ANK.
DR InterPro: IPR000488; Death.
DR Pfam: PF00023; ank. 24.
DR Pfam: PF00531; death. 1.
DR Pfam: PF00791; ank. 1.
DR PRINTS: PR01415; ANKYRIN.
DR SMART: SM00248; ANK; 22.
DR SMART: SM00218; ZUS; 1.
DR PROSITE: PS50088; ANK_REPEAT. 20.
DR PROSITE: PS50297; ANK_REPEAT_REGION. 1.
DR PROSITE: PS50017; DEATH_DOMAIN. 1.
KW ANK repeat; Repeat.
SQ SEQUENCE 1719 AA; 189011 MW; F63465D16D975CBF CRC64;

Query Match 6.2%; Score 545; DB 4; Length 1719;
Best Local Similarity 20.6%; Pred. No. 4,4e-26;
Matches 376; Conservative 270; Mismatches 604; Indels 574; Gaps 71;
21 LKALEKCDVDERNECGOTPLMLAEOGNVEIVELKNGANCLDLDMMWTLISASK 80
94 VELVNYGAVNAQOSKGFPLIYMAQENHLEVKAFLELNGANQNVATEDGFPLAVALO 153

```

```

QY 81 EGH----- 83
Db 154 QGHENYVAHLINVTGKVRALPALHIAARNDORTAAVLNDPNDVLSTGTFPLHIA 213
QY 84 -----IHIVELLKSGASLEHRDMGWTALAMACYGRDVELLSHGANDSVGLQYS 138
Db 214 AHENINLAVOLLNBRASVNETPQNGITPLHIASRGNVIMVLLDREGAQLE-TKTDE 272
QY 139 VYRILMAAGRGHADIHILLONGAKVNSDKVTGTLVMAARGHLECKHILLMAGADV 198
Db 273 LTPHCAARGHVRISIELLDHGAP1QAKTKNGLSPLHMAAQGDHLDVYRLLQYDAED 332
QY 199 -----OEGA-----NSWTALIVAKGTYGYSKEILK 225
Db 333 DITLDHLPPLVAHAGHNRVAVLDDKAKPRSRALNGTPTPLHIACKNHVAVMELLK 392
QY 226 RNRNVNLTDRGTALMISKEGHIEIVODLDAGTYVNPDRSGTTLGAVRGHVEI 285
Db 393 TGAISDAVETSGTLPPLHVASFMGHLPIVKNLDRGASPNVSNKVTETPLHMAARAGHTEV 452
QY 286 VRALLQKYADIDIRGDNKTALYAVEKGNATVNRIOCNPTLETCTGDETELLKATK 345
Db 453 AKTLLOKAKAVNNAKADDOTPLHCAARIGHTNWKLLLENNANPNLATAGHTPLHIAK 512
QY 346 MRNIEVELLDKGAQVSAVDKGDTPPLHVAIRGRSRLAELLRNPKDGLRLYRPNKAG 405
Db 513 EGHVEYTLALLEKEASQACMTKKGTPPLHMAKYGVRAVELLE-RDA-HNNAAG 566
QY 406 E--TPYNDCSHQ-----KSLI-----TOIFGAR----- 427
Db 567 KNGITPLHVAHNNNDIYKLLPRGSGPHSPAMNGYTPPLHIAKONOVEVARSLQYGG 626
QY 428 -----HSPTEGDMIGYDISSALADILSEPTMQP-----PIC-- 462
Db 627 SANAESVQVTPPLHIAQOBGAEMVALLSKQANGNIGKSGTLPPLHVAQOBGVYADV 686
QY 463 -----VGLYAQWMSGK--SEILKLEED-EMKTEAGQOTEPPLFQFQSWL 501
Db 687 LIKHGVAVDATYRNGYTPPLHVAHSHGNIKLVKFLHQAADVNAKTLGY--SPLHQAQO 744
QY 502 --IVFLLLLCGGLGVAFAPVD-----TNLAIALSLALAIYFIFYITVGGRRGE 553
Db 745 GHNDIVYLLKNG-----ASPNEVSSDGTPLAIARLGIYVTDVKVY----- 789
QY 554 SMMAMALSTRLARHIGYLELLFKIMEVNPPELTQTKALPVRFETDNRLS-SVGE 612
Db 790 TDESFYLVSDKHR-MSPRETVDEILDVSEDEBELIS-----FKARRSRDYDEE 840
QY 613 TSAEMIALVSDACEREFGFLATRLFR-----VFTEE-SQGRKKWKTKCCLPSFY 662
Db 841 KELIDFVPKLDQVE-----SPAIPRIPCAMPETVIRSEDEQASKEVEDSLIPS-- 892
QY 663 IFLEFYOCITAGITLAIIFRVDPRKHLTVNALISIASVGLAVLNCRTMWOVLDSLNS 722
Db 893 -----SPATEISDNI-SPVASPVHTGLVSP-----AVDARGGS 925
QY 723 QRRKLSAASKLKLKSEGFMKVLKCEVLMARAKIIDFSTQNOFRLVYIIGDLACBQ 782
Db 926 MRGRNRH-----GLRVVIR-----PRCAAPTITICRLV----- 954
QY 783 DKVLOMLDTPVRLVSKGPFALFASDPHIIKAINONLSVLRDSNNGHDMYRN-IVHL 841
Db 955 --FKOKLSTPPPL-----AEEGLASRIIALGPTGAQFLSPYIVEI 993
QY 842 PVPL-NSRGLSNARKFLYTSATNGDITCSDTTGTQEDTDKRVQNSIGETKIGSTALN 900
Db 994 PHSASHRG--DRELVLVLSSEGSV-----WEHRSRYGESYLDQITLN-GMDEELG 1041
QY 901 RRDYRRROQRTTROMSFDLTKLAVT--EDWESDISPQTRRLNLNITSVYGLLRAN 957
Db 1042 SLELEKKRYCRIT--TDFPLFYVMSRCLQD-YDITGEG-----GSLSKVLPLY 1091
QY 958 QITFNMDRLASWINTLEQWMPYRTSWLILYLEETEGLPDQMTLK-----TWYER 1005

```

```

Db 1092 QATPEPMAVTKRVKALQ-----ACVPDELVTKLLGNQATFSP1VTVPR 1137
QY 1006 ISK-----NIPTTDVEPLEIDODINREYF-----LSSRTPLVLA 1042
Db 1138 RRRFRPPIGLRIPLPSTWDRSGEDDTSLRLCSVIGTQDQAWMEDTGTTKLYA 1197
QY 1043 RDKTFELPCTVNLDBKLEIITADVRAREQINIGGLAVPPLPLHGGPRPSGSOPASV 1102
Db 1198 MECANF--TNV--SARFWLSDCPTAEVNAFALLKEL-----TAPYNAKE 1242
QY 1103 CSSASENPFEPGVSPOPHSSYSGLSGOPHPFYNRAVPATGSSILLSMTVDVCEK 1162
Db 1243 VIFAKMNDPRBG-----MAMNFGD-WHLFRSVMLEMRVSQVPEDEPR----- 1269
QY 1163 LFO-----IEG-----LDONMPOYCTTICKAN-----INGVYL 1191
Db 1270 LEOHNEFVEVARSRDIEVLEGSILPAELSGNLV-----VKRAAORSFHFOSFRENL 1324
QY 1192 SQCNIDELKE-----MAMNFGD-WHLFRSVMLEMRVSQVPEDEPR----- 1233
Db 1325 MGVKVRDSSREPGSSPLRKAMKEDYOHILCHLNTMPACAGSGAEDRRPTPLAL 1384
QY 1234 ---FLNENSSAPVPHGESAR-----SHTLPLTELSSQTPYTLNFSPEELTGLD-- 1283
Db 1385 RYSLISESTPGSLSTGEQEMKAVISEHLGLSMALARE--LOFSVEDINRIYENP 1440
QY 1284 ----EGAPRHSNLSMOSQTRTPPLSLNSODSSEIEKLDKVA-----EKRD 1329
Db 1441 NSLLQSVALLNL-WVIRGQANMENLYTALOSIDREIYVMEGSGRGRNLIKPDHRH 1499
QY 1330 AVREYIAQMSOLEG-----TGSSTISGRSPSTHYVIGOSSGGSIHSTLEOERKE 1382
Db 1500 TBRDYSLSPOQMNQHGROGARITTHSPYVS-QYTERSGORLQDMDQDSIVGLD--AAQ 1556
QY 1383 GELKOE--DGRSFL-----MKRGVIDYSSGVSTNEASPLDPTDEKSDQSGS 1432
Db 1557 GSMOEYVTCGPHSFGTSTMTGELPBGSGOEYKLVASBHTWTEQPEASQADRORR 1616
QY 1433 KLLPGKSSERSPLQOTDLKLLGGGLRYOKPLSDE-----DESGTRVQITPHCSMI 1485
Db 1617 QCGOEVOEAKNTTQVVO-----GNEFONITPEQVTEBOFTDEGN-----I 1660
QY 1486 RTKRLKAK--ORECAS--POEH 1503
Db 1661 VTKIIRKVVROIIDSSADAQEH 1684

```

RESULT 14

070511 PRELIMINARY; PRT; 2622 AA.

AC 070511:

DT 01-AUG-1998 (TREMBLrel. 07, Created)

DT 01-AUG-1999 (TREMBLrel. 11, last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)

DE 270 kDa ANKYRIN G isoform (ANKYRING) (Fragment).

GN ANK3.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RN (1)

RP SEQUENCE FROM N.A.

RX MEDLINE=98417633; PubMed=9744885;

RA Zhang X., Bennett V.,

RT "Restriction of 460/270-kD ankyrin G to axon proximal segments

RL requires multiple ankyrin G-specific domains.";

RL J. Cell Biol. 142:1571-1581(1998).

RN (2)

RP SEQUENCE FROM N.A.

RA Carpenter S.S., Zhang X.;

RL Submitted (Oct-1998) to the EMBL/Genbank/DBJ databases.

RN (3)


```

QY 1345 TGSSTISGRSSPHST----- 1359
DB 1614 APNSTLSRSPPTVTTAGSLKSSITMPPASPKANITWSSSLPFKSIITSAAPLISP 1673
QY 1360 --YTIQSSSGGSHSTLEQERK--EGELKDEGRKSPKMGV--IDYSS----- 1407
DB 1674 LKSVSPKTSADVISTAKAAMASLSPKQMSGAEEVALVNGSVPLKYPSSSALLNG 1733
QY 1408 -----GVSTTEASPL--DPITEDEKSDQSGSKLJPGK--SSERPSTLPQD 1451
DB 1734 CATATLQKIKISTATNAVSIVSAAPDIYEKALSTTTAMPSPSLSYSAAPSAQS-L 1792
QY 1452 KIKGGGLRYQKL-----PSDEDES-----GTGRVQITPHCSKMITKRLKAKOREC 1497
DB 1793 RAPSASALYNLSGSPSVGVTTSVTSIIIVPYSGVNYLAEPALKLPDQNSLITSAAL 1852
QY 1498 AS-----POEH--SAPIRFIKAEYELSDALLDKDSDSGVRSNE-----S 1538
DB 1853 LSPIKTLTETTPROPQPFNTSSP--YKSSLEFLASSAL--KPSVPSSLSSSQELIKDVA 1906
QY 1539 SPNHSILHNEAA--DDSOLEKANLIELEDEGSHGKGMPHSLSG:QDPI-----IARMSI 1590
DB 1907 EKKEDLMKMTALLOQDVPREKFPQDLPREGHIDDEEPKIVKYEKEDLVKSELKQDV 1966
QY 1591 CSEDKKSP-SECSLASPEESW-----PACOKA-----YNLKRTPS 1626
DB 1967 CVESKGPSPKSDKSHSPEDDWTFFSEIIEARQAASHAPSLPERVHGKANLRYID 2026
QY 1627 TYTLNNNTAPTRANDNPEI--EGIRETSQYILRGPSPNPTAVQNEMLK--SMAHK 1680
DB 2027 YLTINDGSSSLTNLKYFEAKREG--EEROKRILKPA--MALQEHKLMPPASMRPS 2080
QY 1681 RRSOR-----SSYTRLSKASELHAA-----SSESTG 1706
DB 2081 TSEKELCKMADSPFGTDALLESPPDESQHDQKSPLSBDSGFTREKTPSAQSAESTG 2139

RESULT 15
ID 012955 PRELIMINARY; PRT; 4377 AA.
AC 012955;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE Ankyrin G.
EN ANK-3.
SS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN STEM.
RX MEDLINE=95138209; PubMed=7836469;
RT Kordeli E., Lambert S., Bennett V.;
RT Ankyrin. A new ankyrin gene with neural-specific isoforms localized
RT at the axonal initial segment and node of Ranvier.";
RL J. Biol. Chem. 270:2352-2359(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN STEM.
RA Carpentier S.S.;
RA Submitted (Aug-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; U13616; AAA64834.1; -.
DR HSSP; P55273; 1B18.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000488; Death.
DR InterPro; IPR000906; ZUS.
DR Pfam; PF00023; ank; 24.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00791; ZUS; 1.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 21.

```

```

DR SMART; SM00005; DEATH; 1.
DR SMART; SM00218; ZUS; 1.
DR PROSITE; PS50088; ANK_REPEAT; 21.
DR PROSITE; PS50297; ANK_REGION; 1.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
DR ANK repeat; Repeat.
SQ SEQUENCE 4377 AA; 480395 MW; F42379E5768B684 CRC64;

Query Match      5 94; Score 526; DB 4; Length 4377;
Best Local Similarity 20.0%; Pred. No. 3.9e-24;
Matches 416; Conservative 266; Mismatches 664; Indels 734; Gaps 80;

QY 14 EENIPALKALEKCDVDENEGQTPIMIAEQWVEYKELKNGANCNLEDLNNMT 73
DB 83 KEHVEYVELLOREANVDATKTKGNALHTASLAGAEVYKVLVTGANVMSQNGFT 142
QY 74 ALISAKSGHIIHYEELIKSGASLEHRDGMGTALMNAKYGRDVELL----- 124
DB 143 PLVMAQENHLEVKFLIDNGASOSLATEDEGFTPLAVALQGHQVVSLLENDTKGVR 202
QY 125 -----SHGANPSYTG----- 135
DB 203 LPAHLIARKDQTAALLLONDNADYESKSGFTPLHIAHYGNINAVATLLNRAAYD 262
QY 136 --QSYVPIITMAAGRGHADIYHLLONGAKVNCSDYGTTPLYWAARKG----- 182
DB 263 FTARNDIPLHVASRGANVVKLLDRGAKIDAKTDGLPLHICGARSQHEVYEMLD 322
QY 183 -----HLECYKHLAMGADVDOGANSMTALIVAKGYTQS 219
DB 323 RAAPLSTKTKNGLSPHATQGDHLCVQLLQHNVPDVTNYTLALHVAACHGHYK 382
QY 220 VKELIKRPNNVNLDDKQNTALMTASK----- 247
DB 383 AKVLIDKANPNNAKALNGFTPLHACKKNRIKYMELLKNGASIOAVTESGLPIHYAAL 442
QY 248 -GHIEIVODLDAGTYVNIIPDRSGDYVLIGAVRGHVEYVALLQKYADIDIRGQDKTA 306
DB 443 MGHVNIYQLMHNHGASPTNTVREGTALHMAARSGAEVYVLYVDAQAYEAKKDDQTP 502
QY 307 LYMAVEKGNATMVDILQ--CNPTD-----EICT 333
DB 503 LHSIRLKGADIVQQLQOGASPNAAVTSYTPHLSAREGHEDVAAFLLDHGASLSITT 562
QY 334 KDGFTPLKATKMNIEVELLDKGAQVSAVDKGGTPPLHVAIRGSRRLAELLNPK 393
DB 563 KKGFTPLHVAAKYKGLVAVANLLLOKASPDAGSGULPLHVAHYDNQVALLL--D 619
QY 394 DGRLLYRNKAGETPYNIDC--SHOKSILTOY--FGAR-----HLSPETDGD 438
DB 620 QGASPHAAKNGYTPPLHIAAKKNQMDIATLILEYGADANAVTRQGIASVYHLAAGEHVDM 679
QY 439 LGYDLISSALDIISEPTMOP-----PICYGLY 466
DB 680 VSLILGRNANVNLNKSGLTPLHLAQEDRYNAVEVLVNOGAHYDAOTKMGYTPHAGCH 739
QY 467 AQWMSGK--SFLIK-KLEDEMKTFAGQGTPLPQFSML--IVFTLLTLCGLGVAFR- 520
DB 740 --YGNIKIVNPLDHSKAVNAKTKNGY--PLHQAQOGGTHIINVLQNN-----ASP 790
QY 521 --VDNLAIASISFLALIVFTIVYFGGRREGESMNAMALSTRIAHIGYELLFK 577
DB 791 ELTVNGNATLGI-----ARLGIISVVD 814
QY 578 LMFV-----NPPE-----LPQOTKALPVRLFTDYNRKLSV--GGET 613
DB 815 LKIYETETMTTITTEKHKMNVPTMNEVELDMSDEVRKANAPBMLSDGEYISDVERGED 874
QY 614 SLA-----EMIAITLSDACEREPFLATRL-----FRVPTESQKKKKKKYCCCL 658
DB 875 AMTGDTKYLGPQDLKELGDSLPAEGYMGFSLGARSASLRSSSDSYTLN--RSSYAR 932
QY 659 PSFVIFLEIVGCIAGITLLAIFRVD--KHLT-----VNLILISIASVGLAFVLC 709

```

```

Db 933 DSMMEELLVPSKEQHLLFTREFDSDSLRHYSMAADTLDNVNLVSSPIHSGFLVSPWDA 992
QY 710 RTWMOVLDSLNSQRRKLHSAASKLHLKSEGFMKVL---KCEVELMARMAKTIDSEFTQ 765
Db 993 RG-----GSMRGSRRHG-----MIIIPRKCCTAP-----TR 1019
QY 766 NOTRLVYIIDGDACEODKVLQMDTVAVLFSKGFIAIFSDHHIITKAINQNLNSTR 825
Db 1020 ITCRLV-----KRRKLAN-----PPPH-----GRRGISSRLV 1047
QY 826 DSNINGHDYMRN-IYHLVPFLNSRGLSNARKFLVTSATNGDITCSDTGTQEDTDRRVSQ 884
Db 1048 EMGPAGAQFGLGVIYIPIHFGSMRG--KERELIVLRSENGE-----TWKEHQPD---SK 1096
QY 885 NSLGEMTKL--GSKTALNRDQTYRRRQOMQRTITROMSFDLTKLLVTEDEWESDISPQWRR 942
Db 1097 NE--DLTEILNGMBEELDPELGRKRICRIITKDF-----POYFAVVS--RIKQ 1142
QY 943 LLNIYSVTGRLLRAN-----QITFMWRLASMINLTEQWPRYTSMLIYLEETGLPDQM 997
Db 1143 ESNQIGPBGILSITTVLVQASPEGALTKRIRVGLO-----AQPVPDEI 1188
QY 998 TLKTM-----YERISKNI-----TTKVEPLLEI 1022
Db 1189 VKKILGNKATFSPIVTEPRRRKFKHPIITMTIWPPEGSEGVNGYKGDTPNLRLLCSI 1248
QY 1023 DG-----DIRNEVFLSSRTPLVARDYKTLPLCTVNDPKRIIADYRAAREQIN 1074
Db 1249 TGGTSPAQWEDI-----TGTPLTLFIKDCVSF---TTNV--SAREWLADCHQVLETVG 1296
QY 1075 IGGIAYPPLPHEGPPRPGSGVSPASVCSASFNGPGGVSPQPHSSYSGLSGPOH 1134
Db 1297 LATQIYREL-----ICV-----PYMAKFV-----VFAKMNDP-- 1323
QY 1135 PTYNRAAVPATGSSLLLSMTVDVYCEKLRQIEGLDNMMPQCTTIKKANINGRVLSQL 1194
Db 1324 -----VSSLRCFCMTDDKVDKTLQEO----- 1346
QY 1195 NIDELAKKEMAMFGMHLFRSMVLEMRSEVQVPEDEPRFLN--ENSSAPVPHGE-----S 1248
Db 1347 NEEV-----ARSKDIEVLEKPIYVDCYGNLAPLFRGGQOVLFN 1386
QY 1249 ARRSHTELP-----LTELSSQTPYTLNFSFELNLTGLDEGAPRHSNLSMOSQTRTPSL 1304
Db 1387 FYSEKENLRFPSIKIRDSQPERCGSLFLKERKTKGLPOTAVCNLNTLPLAKHKTES- 1445
QY 1305 SLSNQSSIEISKLTDKVQAEYRDARE---YIAQMSQLEGGTSSSTISGRSSPHSTYY 1361
Db 1446 ---DQDEIE---KTDROSFAFSLALRRKRYSLTEPGMIRSTGAT---RSLP--TTY- 1492
QY 1362 IGOSSSGSIHSTLEQERKGEGELKQEDGRKSFMLKRGDVIDYSSSGVSTNEASPLDPIT 1421
Db 1493 ---SYRPFSTRPYQSWTAPITVPGPAKS-----GFTSLSSSSSNTPSASPLKSTIW 1541
QY 1422 EDEKSDSGSKLPGKKSSEPSLFOJDLKLKGGLRYQKLPDEDESGTGRVOITPHC 1481
Db 1542 -----SVGRPSPIKSTL-----GASTTSSVKSTSDV 1567
QY 1482 SKMITKRLAKQREKASPOEHSAPITFTIKAEY---LSDALIDK-----KDS 1528
Db 1568 ASPIRSLRT-----MSSPIKTIVVQSOPYNIQVSSGTLARAPAVTEATPLKGL 1614
QY 1529 SDGVRSMESSPNHSLNHEADDQLEKANLIELEDEHSGKRGMPHS----- 1576
Db 1615 ASNSTFSSRTSP-----VTTAGSLERSSITMTTPPASPKSNINNYSSSLPFKSITSA 1668
QY 1577 -----LSGLDPIIARMSICSEDKSPSECSLASSPEESMPA----- 1614
Db 1669 PLISSPLKSVSPVKSRYDVISSAKITMASS---LSSPYKOMPGHAEVALYNGSISPLKY 1725
QY 1615 COKAVNLNRTBSTVTLNHN--TAPITRANO---NEDEIEGIRETSQVI---LRPGSPN 1665

```

Db 1726 ASSSTLLNGCKATATLQEKISSATNSVSVSAATDVEKVFSTTTAMPFSPLRSYVSA 1785

QY 1666 PTAVOENENLKSMHAKRSQSSYTRLSKDSSELHAASSEST 1705

Db 1786 PSAFQSLRTPSAS-----ALYTSL---GSSISATTSVY 1816

Search completed: July 1, 2003, 14:40:08
Job time : 109 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2003, 14:34:04 ; Search time 16 Seconds
(without alignments)
4445.744 Million cell updates/sec

Title: US-10-021-571-4
Perfect score: 8884
Sequence: 1 MSYLIQSIVNYVEENIPA.....LHAASSESTGGERESTL 1715

Scoring table: BLOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	555.5	6.3	1880	1 ANK1_HUMAN	P16157 homo sapien
2	539	6.1	3924	1 ANK2_HUMAN	O01484 homo sapien
3	525	5.9	1862	1 ANK1_MOUSE	O02357 mus musculu
4	521	5.9	4377	1 ANK3_HUMAN	O12955 homo sapien
5	458	5.2	768	1 YB23_HUMAN	O09117 homo sapien
6	438.5	4.9	1059	1 Y379_HUMAN	O15084 homo sapien
7	409.5	4.6	1431	1 DAPK_HUMAN	P53355 homo sapien
8	376.5	4.2	1327	1 TNK1_HUMAN	O05271 homo sapien
9	373.5	4.2	323	1 ANKH_CHRVI	O06527 chromatiu
10	370	4.2	1166	1 ASB2_HUMAN	O09627 homo sapien
11	370	4.2	1166	1 TNK2_HUMAN	O09282 homo sapien
12	351	4.0	741	1 RN5A_HUMAN	O05823 homo sapien
13	351	4.0	692	1 ANR6_HUMAN	O09284 homo sapien
14	348.5	3.9	832	1 ANR3_HUMAN	P57078 homo sapien
15	348	3.9	735	1 RN5A_MOUSE	O05921 mus musculu
16	330.5	3.7	518	1 ASB3_HUMAN	O09575 homo sapien
17	328	3.7	1401	1 LAT1_LATMA	P23631 latrodectus
18	324.5	3.7	525	1 ASB3_MOUSE	O09472 mus musculu
19	318.5	3.6	747	1 V222_FOPV	O09513 fowlpox vir
20	318	3.6	583	1 AS15_MOUSE	O09513 fowlpox vir
21	309	3.5	596	1 V244_FOPV	O09517 fowlpox vir
22	307	3.5	668	1 V244_FOPV	O09517 fowlpox vir
23	304.5	3.4	542	1 V155_FOPV	O09567 fowlpox vir
24	302	3.4	436	1 V245_FOPV	O09445 fowlpox vir
25	284	3.2	1083	1 Y1L2_YEAST	P40460 saccharomyc
26	278	3.1	525	1 V228_FOPV	O09507 fowlpox vir
27	275	3.1	592	1 V246_FOPV	O09544 fowlpox vir
28	267	3.0	603	1 V162_FOPV	O09569 fowlpox vir
29	264.5	3.0	656	1 FEM1_CAEEL	P17221 caenorhabdi
30	259	2.9	433	1 AS14_MOUSE	O08587 mus musculu
31	254.5	2.9	434	1 AS15_HUMAN	O09441 homo sapien
32	252	2.8	642	1 YAZA_SCHPO	O09701 schizosacch
33	242.5	2.7	429	1 AS10_HUMAN	O08wx13 homo sapien

34	240.5	2.7	278	1 AS13_HUMAN	O08wx3 homo sapien
35	240.5	2.7	776	1 ANR5_HUMAN	O09u02 homo sapien
36	239	2.7	434	1 V023_FOPV	O09158 fowlpox vir
37	235	2.6	333	1 ANR2_HUMAN	O09421 homo sapien
38	233.5	2.6	2703	1 NOTC_DROME	O07207 drosophila
39	232.5	2.6	578	1 V022_FOPV	O09159 fowlpox vir
40	231.5	2.6	328	1 ANR2_MOUSE	O09406 mus musculu
41	231.5	2.6	437	1 V014_FOPV	O09517 fowlpox vir
42	231.5	2.6	461	1 V218_FOPV	O09517 fowlpox vir
43	230	2.6	226	1 PSDA_HUMAN	O075832 homo sapien
44	229.5	2.6	1066	1 NDC2_NEUCR	O01317 neurospora
45	229	2.6	231	1 PSDA_RAT	O0922x3 rattus norv

ALIGNMENTS

RESULT 1
ID ANK1_HUMAN STANDARD: PRT: 1880 AA.

AC P16157;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ankyrin 1 (Erythrocyte ankyrin) (Ankyrin R).
GN ANK1 OR ANK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND VARIANTS.
RC TISSUE-Hematopoietic;
RX MEDLINE=90158830; PubMed=2137557;
RA Lux S.E., John K.M., Bennett V.;
RT "Analysis of cDNA for human erythrocyte ankyrin indicates a repeated structure with homology to tissue-differentiation and cell-cycle control proteins.";
RT Nature 344:36-42(1990).
RL [2]
RN SEQUENCE FROM N.A.
RP MEDLINE=90175370; PubMed=1689849;
RX Lambert S., Yu H., Prohal J.T., Lawler J., Ruff P., Speicher D., Cheung M.C., Kan Y.W., Palek J.;
RT "cDNA sequence for human erythrocyte ankyrin.";
RT Proc. Natl. Acad. Sci. U.S.A. 87:1730-1734(1990).
RN [3]
RX VARIANT HS ILE-462.
RX MEDLINE=96225450; PubMed=8640229;
RA Eber S.W., Gonzalez J.M., Lux M.L., Scarpa A.L., Tse W.T., Dornwell M., Herberts J., Kugler W., Oezcan R., Pekrun A., Gallagher P.G., Schroeter W., Forget B.G., Lux S.E.;
RT "Ankyrin-1 mutations are a major cause of dominant and recessive hereditary spherocytosis.";
RT Nat. Genet. 13:214-218(1996).
CC -1- FUNCTION: ATTACH INTEGRAL MEMBRANE PROTEINS TO CYTOSKELETAL ELEMENTS. BIND TO THE ERYTHROCYTE MEMBRANE PROTEIN BAND 4.2. TO NA-K ATPASE, TO THE LYMPHOCYTE MEMBRANE PROTEIN G85, AND TO THE CYTOSKELETAL PROTEINS FODRIN, TUBULIN, VIMENTIN AND DESMIN.
CC ERYTHROCYTE ANKRYNS ALSO LINK SPECTRIN (BETA CHAIN) TO THE CYTOPLASMIC DOMAIN OF THE ERYTHROCYTES ANION EXCHANGE PROTEIN; THEY RETAIN MOST OR ALL OF THESE BINDING FUNCTIONS.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC SURFACE OF ERYTHROCYTIC PLASMA MEMBRANE.
CC -1- ALTERNATIVE PRODUCTS: At least 3 isoforms; 1/2.1 (shown here), 2/2.2 and 3, are produced by alternative splicing.
CC -1- PTM: REGULATED BY PHOSPHORYLATION.
CC -1- PTM: ACYLATED BY PALMITIC ACID GROUP(S).
CC -1- DISEASE: Defects in ANK1 are the cause of dominant and recessive hereditary spherocytosis (HS).
CC -1- SIMILARITY: CONTAINS 23 ANK REPEATS.
CC -1- SIMILARITY: CONTAINS 1 DEATH DOMAIN.

[illegible][illegible]


```

QY 835 MRNIVHLPVFLNSRGLSNARKFLVTSATNGDVPSCSDTIGDADDRVSONSLGEMTK-- 892
DB 1135 -----GGVLSSTVPVQVAVPPEALFKRI-----RVGLQ-----AQMHSSELYVKKI 1176
QY 893 LGSKTALNRDYY--RRQMORTITROMSEDLTKLLVTEDEWSDISPOTMRRLINIVSYT 950
DB 1177 LGNKATFSPVLTLEPRRRKFKHPTITWIPVPAKSSDVMNLNGFGDAP-TLRLL--CSIT 1232
QY 951 GRLLRANQISFNNDRLASWINLT--EQMPYRTS-----WLLIYLETGTIGDQMTLKI- 1001
DB 1233 GGTTPA-----OMEDITGTPTPLTFVNECVSFTTNVSARPMLI--DCRQIOESYVFAQ 1283
QY 1002 IYERI-----SKNIPTTKDVEPL-----EIDGDIRNFVEFSRTPVAVAD 1044
DB 1284 VYREILICVPYMAKFVFAKSHDPIEARKLPCFCTDKVKTKLEQCNF--AEVARSKD 1339
QY 1045 VKV-----FLPCTVNLDP-----KLREIAD--VRAAR 1070
DB 1340 VEVLGKRPVYVDFCGMLVPLTKSGQHIFSFAPKERNRLPLFKVYKVDITQEPGRLSPMK 1399
QY 1071 EQISIGLAVP-----PLPLHGPFRAPSGYSGPPVCSSTFNGPAGVSPQPS 1123
DB 1400 EPRSTGLVHQAICNINILPTLY--TKESSDQDEEEDIMTSEKN-----DETBS 1448
QY 1124 SYSGMTGPOHPFYNNGSGPAPGVVULNSLNDVACEKLQIEGDMLEPOYCTTIKK 1183
DB 1449 TETSVL--KSH-----LVNEVPYASPDLSSEKMKODLIK--TAILT 1489
QY 1184 ANINGVLAQCINIDELKEMNMFGMHLFRSTVLEMRNASHVDEDPFLISE--SSSGP 1242
DB 1490 TDVSDA--GSIKYKEIYKAAEEBEPGE--PFEIVER-----VKEDELEKVEILIRSGT 1537
QY 1243 APGGEFARRASHHELHTELSQTPYTLNFSFEELTIGLDEGAPRHSNLSMOSQRRRP 1302
DB 1538 CTDESSVSSRSERGLVE--EEMVIVSDEIEEAR--QAPLEITEYPCVEVRIDK 1590
QY 1303 SLSSLSNODSIEISKLTQKVAQAEYDAYREYIAOMSOLEGGPGSTTISGRSSPHSTYMA 1362
DB 1591 EIKGVKEKSTGLVNLITDILNT--CYPLPREQLQYQDAGKCEALAVGRSS----- 1642
QY 1363 GQSSSGGSIHNSLEQEKGDSEP-----KPDGGRKSLFMKGDVIVYSSSGVSTIN 1412
DB 1643 -----EKEGKIDIPDETOSTQOKHKKPGLGIKKPVRK----- 1674
QY 1413 DASPLDPIPEDEKSSQSSGSKLLPGKSSRSLSLQTDLKLKSSGLRYOKLPSEDESGT 1472
DB 1675 -----LEKOKQKEEGLQASAEKAEKELKGSSESLGE-----DGLAPEPLPVKATS-- 1722
QY 1473 EESDNTPLTKDDDKRAEGKVERVPKSP-----HSAEPIRTFIRKAKEYLSDALDKDS 1527
DB 1723 -----PLIEETPIGSKDKVKALQKRVDEQGRKSLPIR--VKKEDEVPKTTIRPH 1774
QY 1528 SDSGVNSF-----SSPNHSLHNEVADSDQLEKANLIELEDSSHSGKRGIPHSLSGLD 1581
DB 1775 AASPILKSEHRAHAGSPKTERHSTLSSAKTERHPVVS-----PSSKTEKHS 1822
QY 1582 PIARMSICSEDKKSPSECLIASPENNPAQOKAVNLMRTSTYTLNNSKAPARANG 1641
DB 1823 PV--SPSAKTERHSPSSS--SKTEKHSV-----SPSTTERHSPVSSKTER 1867
QY 1642 NPEDEMGIRETSQVILLRPPSSPNPTIIONENLK-SMTHKRSQSSYTRLSKDP 1694
DB 1868 H-----PPVSPSGTKDRKRPVSPSGRTEKHPVSPGRTKRLPVSPSGRDKHOP 1917

```

```

DE Ankyrin 1 (Erythrocyte ankyrin).
GN ANK1 OR ANK-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN NCBI_TaxID=10090;
RP SEQUENCE FROM N.A.
RC TISSUE=Erythrocyte;
RX MEDLINE=92345717; PubMed=1386265;
RA White R.A., Birkenmeier C.S., Peters L.L., Barker J.E., Lux S.E.;
RT "Murine erythrocyte ankyrin cDNA: highly conserved regions of the
RT regulatory domain.";
RL Mamm. Genome 3:281-285(1992).
CC - FUNCTION: ATTACH INTEGRAL MEMBRANE PROTEINS TO CYTOSKELETAL
CC ELEMENTS. BIND TO THE ERYTHROCYTE MEMBRANE PROTEIN GP85, AND TO THE
CC NA-K ATPASE, TO THE LYMPOCYTE MEMBRANE PROTEIN GP85, AND TO THE
CC CYTOSKELETAL PROTEINS FODRIN, TUBULIN, VIMENTIN AND DESMIN.
CC ERYTHROCYTE ANKYRINS ALSO LINK SPECTRIN (BETA CHAIN) TO THE
CC CYTOPLASMIC DOMAIN OF THE ERYTHROCYTES ANION EXCHANGE PROTEIN;
CC THEY RETAIN MOST OR ALL OF THESE BINDING FUNCTIONS.
CC - SUBCELLULAR LOCATION: CYTOPLASMIC SURFACE OF ERYTHROCYTIC
CC PLASMA MEMBRANE.
CC - PTM: REGULATED BY PHOSPHORYLATION (BY SIMILARITY).
CC - PTM: ACYLATED BY PALMITIC ACID GROUP(S) (BY SIMILARITY).
CC - SIMILARITY: CONTAINS 23 ANK REPEATS.
CC - SIMILARITY: CONTAINS 1 DEATH DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: M84756; AAA37236.1; -.
DR HSSP: Q00420; IAWC.
DR MGD: MG1:88024; ANK1.
DR InterPro: IPR002110; ANK.
DR InterPro: IPR000488; Death.
DR InterPro: IPR000906; Z05.
DR Pfam: PF00023; ank; 24.
DR Pfam: PF00531; death; 1.
DR Pfam: PF00791; Z05; 1.
DR PRINTS: PR01415; ANKYRIN.
DR SMART: SM00248; ANK; 22.
DR SMART: SM00005; DEATH; 1.
DR PROSITE: PS50088; ANK_REPEAT; 20.
DR PROSITE: PS50297; ANK_REGION; 1.
DR PROSITE: PS50017; DEATH_DOMAIN; 1.
KW Cytoskeleton; Repeat; ANK repeat; phosphorylation; lipoprotein.
FT DOMAIN 1 827
FT 828 1386
FT 1387 1862
FT DOMAIN
FT REPEAT 40 69
FT REPEAT 73 102
FT REPEAT 106 135
FT REPEAT 139 168
FT REPEAT 170 197
FT REPEAT 201 230
FT REPEAT 234 263
FT REPEAT 267 296
FT REPEAT 300 329
FT REPEAT 333 362
FT REPEAT 366 395
FT REPEAT 399 428
FT REPEAT 432 461
FT ANK 1.
FT ANK 2.
FT ANK 3.
FT ANK 4.
FT ANK 5.
FT ANK 6.
FT ANK 7.
FT ANK 8.
FT ANK 9.
FT ANK 10.
FT ANK 11.
FT ANK 12.
FT ANK 13.

```


Db 1730 GEPALOEBOVLV---STREHVOGPETGSPRAGKPELUMAFESAFS 1776

RESULT 4
ID ANK3_HUMAN STANDARD; PRT: 4377 AA.
AC Q12955;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ankyrin 3 (ANK-3) (Ankyrin G).
GN ANK3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
NC TISSUE=Brain stem;
RX MEDLINE=95138209; PubMed=7836469;
RA Kordell E., Lambert S., Bennett V.;
RT Ankyrin. A new ankyrin gene with neural-specific isoforms localized
RT at the axonal initial segment and node of Ranvier.";
RL J. Biol. Chem. 270:2352-2359(1995).
CC -1- FUNCTION: Membrane-cytoskeleton linker.
CC -1- ALTERNATIVE PRODUCTS: A number of isoforms are produced by
CC alternative splicing.
CC -1- TISSUE SPECIFICITY: Expressed in brain and other tissues.
CC -1- SIMILARITY: CONTAINS 23 ANK REPEATS.
CC -1- SIMILARITY: CONTAINS 1 DEATH DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

DR EMBL: U13616; AAA64834.1; -
DR HSSP: P5273; IB18.
DR Genew: HGNC:494; ANK3.
DR MIM: 600465; -
DR InterPro: IPR002110; ANK
DR InterPro: IPR000488; Death.
DR InterPro: IPR000906; Z05.
DR Pfam: PF00023; ank; 24.
DR Pfam: PF00531; death; 1.
DR Pfam: PF00791; Z05; 1.
DR PRINTS: PR01415; ANKYRIN.
DR SMART: SM00248; ANK; 21.
DR SMART: SM00005; DEATH; 1.
DR SMART: SM00218; Z05; 1.
DR PROSITE: PS50088; ANK_REPEAT; 21.
DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE: PS50017; DEATH_DOMAIN; 1.
KW Cytoskeleton; Alternative splicing; Repeat; ANK repeat.
FT REPEAT 73 102 ANK 1.
FT REPEAT 106 135 ANK 2.
FT REPEAT 139 168 ANK 3.
FT REPEAT 172 201 ANK 4.
FT REPEAT 203 230 ANK 5.
FT REPEAT 234 263 ANK 6.
FT REPEAT 267 296 ANK 7.
FT REPEAT 300 329 ANK 8.
FT REPEAT 333 362 ANK 9.
FT REPEAT 366 395 ANK 10.
FT REPEAT 399 428 ANK 11.
FT REPEAT 432 461 ANK 12.
FT REPEAT 465 494 ANK 13.
FT REPEAT 498 527 ANK 14.
FT REPEAT 531 560 ANK 15.

FT REPEAT 564 593 ANK 16.
FT REPEAT 597 626 ANK 17.
FT REPEAT 630 659 ANK 18.
FT REPEAT 663 692 ANK 19.
FT REPEAT 696 725 ANK 20.
FT REPEAT 729 758 ANK 21.
FT REPEAT 762 791 ANK 22.
FT REPEAT 795 825 ANK 23.
FT DOMAIN 1519 1898 SER-RICH.
FT 4090 4174 DEATH.
SQ SEQUENCE 4377 AA; 480399 MW; F42379E5768B684 CRC64;
Query Match 5.9%; Score 521; DB 1; Length 4377;
Best Local Similarity 20.6%; Pred. No. 3,86-19;
Matches 427; Conservative 284; Mismatches 695; Indels 668; Gaps 87;
QY 14 EEENIPALKALKCKDVERNECGOTPLMAEÖGNELEYKELIKNGANCNEEDDNMT 73
DB 83 KEHVEVYSELLÖREAVNDATKKGNTALHIASIAGÖEVVKVLTNGAVNNAÖSÖNGFT 142
QY 74 ALIASKRGHYHYEELLCKGVNLEHDMGWTALMAACYKGRDYELL----- 124
DB 143 PLVMAÖENHLEVVKFLIDNGASÖSLATEDGFTPLAVALÖÖGHÖVSLLENDTKGKVR 202
QY 125 -----SHGANPSVTGL----- 135
DB 203 LPALHIAARKDQTKAALLÖDNNDADYESGFTPLHIAHYGINIVATLLNRAAYD 262
QY 136 ---QYVYPTIWAAGRGHADVILLÖNGAKVNCSDRYGTPLVMAARKG----- 182
DB 263 FTARNIDPLHVAKRGANNAVKKLLDRGAKIDAKTRDGLPLHCGARSGHEOVEMILD 322
QY 183 -----HLECYKHLANGADVÖDÖGANSMTALVAVKGYÖS 219
DB 333 RAAPILSTFKKGLSPHMAÖTGDHNCVÖLLÖHNVPDVTYDLYALHVAHCHGYKV 382
QY 220 VKEILKRNPNVNLÖDKÖGTALMASKE----- 247
DB 383 AKVLLDKKANNAALNÖFTPLHIAKKNRKIKVEMELLKHGASIQÖATESGLPIHYAA 442
QY 248 -GHTEIVÖDLDAGTYVNIÖPDSGDTVLIGAVRGHVEYVALLÖKADIDIRÖÖDKTA 306
DB 443 MGHVNIYSÖLHNHGSPNTTVRGETALHMAARSGÖEVVYVÖDÖGÖVAKAKDÖTP 502
QY 307 LYMAVCKGNATMVNDILO--CNPD-----EICT 333
DB 503 LHSIRLGRKADIVÖÖLÖÖGASPNMAÖTSGTPLHLSARSGHEDVAAFLDGHGASLSITT 562
QY 334 KDGEPLIKATKMRNIEVEELLÖDGAÖVSAVDRKGDPTPLHIAIRGRSRKLAELLRNPK 393
DB 563 KKGFTPLHVAKYGKLEVANTLLÖKSASPDAGSGGLTPLHVAHYNÖKÖVALLLL--D 619
QY 394 DÖRLLYRNKAGETPYNIDC--SHÖKSITÖI--FGAR-----HLSPEÖDGM 438
DB 620 ÖGASPHAAKNGYPLHIAARKNÖMDIATLLEYGADANAVTROGIASVHLAÖEGHVDM 679
QY 439 LGYDYSALADIÖSEPTMÖP-----PICYGLY 466
DB 680 VSLLGRNAVNLSSKSGTPLHLAÖEDRVNAVEVLVNCÖAHYDAÖTKMGYTPLVHÖCH 739
QY 467 AÖWGSÖG--SFLÖK--KLDEÖKTFAGÖÖIEBPLÖFSWLVPLTLLÖGÖ--LGLFÖF 519
DB 740 --YGNIKIVNLLÖSÖKAVNAKTKNG--YTPHLÖGÖAÖ-----ÖGHÖHINVLÖN 785
QY 520 TVHFN-----LGIÖVSLSFÖLALYIÖFYIÖYFGRGSGESMNAVSTLÖAH-I 569
DB 786 NASPNELTVNÖNTALGIARLGYISVDTLKIÖV-----EETMTTIVTERHKM 834
QY 570 GYELLLKIMFVNPEL-----PEÖTKALPVRFLÖTDYNRLSV--GÖETSIA----- 616
DB 835 NVPEÖTNEVYLDMSDÖVKANAPE-----MÖSDÖETISYÖEGÖDAMÖDÖTKYL 884
QY 617 --EMÖTÖLSDACERÖFGFLÖATRL-----FRVÖKTED--TÖGKKRKKÖTCLPSFVLÖF 666

```

Db      885 GPQDLKELGDDSLPAEGWGFSLGARASLSFSSDRSYTLNRRSYAR-----DSMIDEL 940
Qy      667 IIGCIISGITTLLAIRVPD-KHLT-----VNAVLSIASVGLAVLNCRTWMOYLD 717
Db      941 IVPSEKQHLTFTREDSDSLRHYSWAADTLDNVNLVSSPIHSGFLVSPVARG----- 994
Qy      718 SLINSQRLHNAASKLHKLSEGFMYL-----KCEVELMARMAKTIDSPONQOTRIYVI 773
Db      995 ---GSMRGRHHG-----MRLIIPKCTAP-----TRICRLV-- 1025
Qy      774 IDGLACQDQVLOMLDTVRLVLESKPPIALFASDPHILIKAINONLNSVLRDSNINGHD 833
Db      1026 -----KRKKLAN-----PPPH-----GERGISSRLVEMGPAGAQ 1055
Qy      834 YMRN-IYHLVPLNRGLSNARKFLYTSATNGDVCSTTGIOEDAPRVSQNSIGETK 892
Db      1036 FLGPVIVELPHFGMRG--KERELIVLSENGE-----TWKEHQFD--SKNE--DLTE 1102
Qy      893 L--GSKTALNRDRTARRROMORTTIRKMSFDLTKLLVTEDEWESDISPQTMRLNIVSVT 950
Db      1103 LLNGDEDELDESEELGKRIRIITKD-----PQYFAYVS--RIKQESNOIGPE 1150
Qy      931 GRLRAN-----QISFNWDLASWINTLEQWPRTSWLILYLETEGIPDMQTLKTYER 1005
Db      1151 GGLISTTVPLVQASFPFGALTKRIRVGLQ-----AQPVDEIVKILGNK 1196
Qy      1006 ISKNPTKQVPELLEIDGIRNFEVLISRTPVIVARDVAVFLPCYVNLDPKRLREIAD 1065
Db      1197 AT-----FSPITVTPRRKRFRKPTMTTPV----- 1222
Qy      1066 VRAAREQISIGLAVPLPLHEGPRAPSGYSOP--PSVCSSTFNGFAGVVSPOPHS 1123
Db      1223 -----PPSGEGVSNYKGDTPPLRLCISIT-----GGRSP-- 1254
Qy      1124 SYISGMPQPHRYNRGSGPARGPVLLN-----SLAND-----AVCEKLQIGLDQSM 1173
Db      1255 AQMEDITQT-----PLTFKIDCVSFTTNVSARFMDLCHQVLETGLATOL 1301
Qy      1174 -----LPOYCTTIKANINGRVLACQNI-----DELKK--EMMANFGDMLFSTYLEM 1220
Db      1302 YRELICVYMAKFVFAKANDPVESSLRCPCMTDKDKVDTLEQDENE-----VA 1332
Qy      1221 RNASHVVPEDPRELSESSSGPARHGEPARA-----SHNELPHT--ELSSQTPYT 1269
Db      1353 RSKDLEVLGKPIYV-DCYGNMLAPLTKGQQLVFNFFYSFKRNRLPFSIKIRDTQBPGR 1411
Qy      1270 LNFSEELNTGLDEGARHNSLSQSGTRTTPSLSLNSQSSIEISKLTDKYQAYRD 1329
Db      1412 LSLFKERTTKGLPQTAVCNINITLPAHKKETES-----DODDEIE--KIDRQSPASL 1463
Qy      1330 AYRE---YIAQMSOLEGGPGSTTISGRSPHSY-----YMQSSSGGSIHNSLE 1376
Db      1464 ALRKRYSLTERPGMIERSTGAT-----KSLP--TTSYAPFESTRYQSMTTAPITV----- 1513
Qy      1377 QEKGDSEPKPDGRKSFMLKRGVLDYSSSVSTNDASPLDPT---TEDEKSDSGS 1432
Db      1514 -----PGPAKSGF-----TSLSSSSMTPASPLKSIWSTPSPKISTGLAS 1556
Qy      1433 KLLPGKSSERSLSLFQDILKKGSLR---YQKLPSDEDEGT-----EESDNTPL-- 1480
Db      1557 TTSSVKSISDVASPIRS-LRTMSSPIKTIVVSQSPYNQVSSGTLARAPAVTEAPLKLGLA 1615
Qy      1481 -----LKDDKRAEKGVER-----VPSPE-----HSAEDIRFI--KAKYLSDAL 1521
Db      1616 SNTFSSRTSPVYTAGSLERSSTMTTPPASPKSNIMYSSSLPFKSIITSAAPLSSPL 1675
Qy      1522 -----LDKQSSDSGVRSSESSPNHSL--HNEVA-----DDSOLEKAN--LLE-- 1560
Db      1676 KSVSPYKSRVDYISSAKITWASSLSIPVKQMPGHABALVNGSISPLKASSSTLNGC 1735
Qy      1561 -----LEDDSHSGKRGIPHLSGLDPIIARMSICSEDKKSPSCSLASSPENWPAQ 1615

```

```

Db      1736 KATATIQEKISSATNSVSSVSAATDVEKVFSTTAMPSPRLR-SYVSAP----- 1786
Qy      1616 KAYLNFTPESTVTLLNNNSAPANRANQNFEMEGIRETSOYILRPSSS-----PNPTTON 1670
Db      1787 SAFQSLTTPSASALYSLSGSSISATIS-----SYTSSITVPVSVNVLPALAKTL 1839
Qy      1671 ENLKMTHKRSQSSS--YTRLSKDPPELHAASS 1702
Db      1840 PDSNSTKSAALLSPKITLTETHPOPHSRYS 1873

```

```

RESULT 5
YB23_HUMAN
ID YB23_HUMAN STANDARD; PRT; 768 AA.
AC Q9ULJ7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, last sequence update)
DT 15-JUN-2002 (Rel. 41, last annotation update)
DE Hypothetical protein KIAA1223 (Fragment).
GN KIAA1223
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=20039619; PubMed=10574462;
RA Nagase T., Ishikawa K.-I., Kikuno R., Hiroseawa M., Nomura N.,
RA Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XV.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 6:337-345(1999).
CC -1 SIMILARITY: CONTAINS AT LEAST 14 ANK REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: AB033049; BAA86537.1; -
DR HSSP: PA2771; IDC2
DR InterPro: IPR002110; ANK.
DR SMART: SM00248; ANK; 14.
DR PROSITE: PS50088; ANK_REPEAT; 13.
DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
KW Hypothetical protein; Repeat; ANK repeat.
FT NON_TER 1
FT REPEAT 1
FT REPEAT 15
FT REPEAT 48
FT REPEAT 86
FT REPEAT 119
FT REPEAT 152
FT REPEAT 185
FT REPEAT 218
FT REPEAT 251
FT REPEAT 284
FT REPEAT 317
FT REPEAT 346
FT REPEAT 350
FT REPEAT 383
FT REPEAT 412
FT REPEAT 416
FT SEQUENCE 768 AA; 82819 MW; 2913B69BE2DFE06D CRC64;
Query Match 5.2%; Score 458; DB 1; Length 768;
Best Local Similarity 27.9%; Pred. No. 5.8e-17;
Matches 122; Conservative 78; Mismatches 163; Indels 74; Gaps 4;

```


Db 545 TMDLSDSPNRATIS-PLHLAAVGHQALEVYOSLIDLVYRNSGRTPLDLAAFKHVE 603
 QY 504 FETLLGGIGLL 516
 Db 604 CVDVLINOGASTL 616

RESULT 7
 DAPK_HUMAN
 ID DAPK_HUMAN STANDARD; PRT: 1431 AA.
 AC P53355;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Death-associated protein kinase 1 (EC 2.7.1.-) (DAP kinase 1).
 GN DAPK1 OR DAPK.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=95129831; PubMed=7828849;
 RA Deiss L.P., Feinstein E., Berissi H., Cohen O., Kimchi A.;
 RT "Identification of a novel serine/threonine kinase and a novel 15-kD
 RT protein as potential mediators of the gamma interferon-induced cell
 RT death";
 RL Genes Dev. 9:15-30(1995).
 RP REVISIONS TO 164-171.
 RA Feinstein E.;
 RL Submitted (Apr-1997) to the EMBL/GenBank/DBJ databases.
 CC -! FUNCTION: INVOLVED IN MEDIATING INTERFERON-GAMMA-INDUCED CELL
 CC DEATH.
 CC -! PTM: AUTOPHOSPHORYLATED.
 CC -! SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC -! SIMILARITY: CONTAINS 10 ANK REPEATS.
 CC -! SIMILARITY: CONTAINS 1 DEATH DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: X76104; CAA53712.1; -;
 DR HSSP: Q63450; IAO6.
 DR GeneW; HGNC:2674; DAPK1.
 DR MIM: 600831; -;
 DR InterPro: IPR002110; ANK.
 DR InterPro: IPR000488; Death.
 DR InterPro: IPR000719; Euk_kinase.
 DR InterPro: IPR002290; Ser_thr_kinase.
 DR Pfam: PF00023; ank; 8.
 DR Pfam: PF00069; kinase; 1.
 DR Pfam: PF00531; death; 1.
 DR ProDom: PD000001; Euk_kinase; 1.
 DR SMART: SM00248; ANK; 7.
 DR SMART: SM00005; DEATH; 1.
 DR SMART: SM00220; S_TKC; 1.
 DR PROSITE: PS50088; ANK_REPEAT; 6.
 DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
 DR PROSITE: PS50017; DEATH_DOMAIN; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 DR Transferrase; Serine/threonine-protein kinase; Calmodulin-binding;
 DR Phosphorylation; ATP-binding; Repeat; ANK repeat; Apoptosis.
 FT DOMAIN 13 266
 FT DOMAIN 334
 FT REPEAT 378 407 ANK 1.

FT REPEAT 411 440 ANK 2.
 FT REPEAT 444 473 ANK 3.
 FT REPEAT 478 507 ANK 4.
 FT REPEAT 511 540 ANK 5.
 FT REPEAT 544 573 ANK 6.
 FT REPEAT 577 606 ANK 7.
 FT REPEAT 610 639 ANK 8.
 FT REPEAT 676 705 ANK 9.
 FT REPEAT 716 745 ANK 10.
 FT NP_BIND 1313 1397 DEATH.
 FT NP_BIND 19 27 ATP (BY SIMILARITY).
 FT BINDING 42 42 ATP (BY SIMILARITY).
 FT ACT_SITE 139 139 BY SIMILARITY.
 FT MUTAGEN 42 42 K->A: LOSS OF ACTIVITY.
 SQ SEQUENCE 1431 AA; 160017 MW; 9EE84811004A155B CRC64;

Query Match 4.6%; Score 409.5; DB 1; Length 1431;
 Best Local Similarity 22.0%; Pred. No. 5,9e-14;
 Matches 225; Conservative 149; Mismatches 338; Indels 309; Gaps 36;

QY 5 ISOSVINYVEENIPALKLECK--DYDERNECGQTPLMTAEQGNLEIKELIKNGA 62
 Db 344 VMKAIHAIINDVNPGLQHLGSLSNYDVNPKNHGTPLLIAAGCGNIQLILIKRGS 403
 QY 63 NCMLEDDIMWTALISAKGKHVIVEELKCGVNLHHRMGGMWTALMMACYKGRDYVEL 122
 Db 404 RIDVQDKGSSNAVYMAARKGHDTLKFLENKCPIDYKKSSEMALHYAARYGHADVAOV 463
 QY 123 LLSHGANPSVYGLQSVYPIIMAAGRGADIYHLLONGAKVNSDKYGTPLVMAARKG 182
 Db 464 TCASAQIDPISRTKEBETPLHCAAHGYSAVKALCEACGNVNIKREGETPLTASARG 523
 QY 183 HLECVHLLAMGADVDEGANSWTALIVAKGYQSVYELLKRNPNVLTDKDNTALM 242
 Db 524 YHDIYVCLAEHGD-----LNACDKGHIALH 550
 QY 243 IASKKEHTIEVDLDAGTYVNPDRSGDTVLIGAVRGHVEIVRALLOKYDIDIRGOD 302
 Db 551 LAYRRCQMEVYKTLISQGFVDYQDRHGTPPLHVACKDSNMIIYALCE----- 599
 QY 303 NKTAIYAVKGNATVRVLILOCPNPTETCTGDETPLIKATKRNIEYVELLDKAKV 362
 Db 600 -----ANCNLD--ISNRYGTRPPLAANNGLDVRVLCIMGASV 637
 QY 363 SAVDKKGPPLHLAIGRSKLAELLRNPKGRLLYRNKAGETPPYNDCHOKSILMQ 422
 Db 638 EALTTCGKTAEDLAREQEHVAGILARKD-----THKELFTQ 678
 QY 423 IFGARHLSPTETDGMGLGYDLYSSALADILSEPTMQPPICVGLYAMQSGKSFLLKLED 482
 Db 679 -----LRPTQ-----NQPRIKLKLFGHSSGSKTTLVESLK- 709
 QY 483 EMTFRAGQIETPLPQFSMLIVLTLLCGGLGLAFYVHPMG----- 526
 Db 710 -----CGILSFFR-RRRPLSSTNSRPPPSPLAKP 741
 QY 527 -IAVSLFLALYIFETIVYFCGRREGESMNAWVLSRLAHIGYELLKLIMVNPPE 585
 Db 742 TVSVSTNNL-----YPCGENSVRSNMFEGGLK--GMLE-----YFVAPTH 783
 QY 586 LP-----EQTRAL-----PVRFLEFDYNRLSVSGESTLAMI 619
 Db 784 HPHCSADDSQTAIDQNAVINGVDVSWERSGNPVYFCYCD---FAANDPISIHVV 840
 QY 620 AILSDACEBEG--FLATRLFRVKTEDT---QSKKKWKTKCCIPSVIFLFIIGCIIS 673
 Db 841 FSLSEPEYETQINPVIFWLSFLKSLVVEEPLAFGGLK-----NPLQVVL----- 885
 QY 674 GTTLAIFRVDKHLTVNAVILSIASVYGLFVLMCRFMQVQVLDLSLQSKRLHN--A 730
 Db 886 -----VATRADINVRPRAGGEGYDKD-----SLKEITNREGNDLHT 925
 QY 731 ASKLHLKNSGEF---MKVLKCEVELMARMAKTTIDSFQNOTRLVVIIDGADCEQDKVL 786

QY 212 -----VKG-GYTQSVKE----- 222
 DB 323 KERLAEPFGHSLQAAAREADYTRIKKHLSELVNPNKHQTHETALHCAASPYPKRKOI 382
 QY 223 ---ILKRNPNVLTQKGTALMIASKEGTEIVODLLAGTYVINIDRSSGDTVLIGAVR 279
 DB 383 CELLRRKANINKEKTEFLEPLHVAASEKANDVEVAVVHEAKVNALDNLGQSLHRAAY 442
 QY 280 GGHVEIVRALLOKAYADIDIRGONKATLVAVKE-----GNATMVRDILQCNP- 327
 DB 443 CGHLQFCRLLSYGCOPNIISLOGFPLQMGNNVOQLLOEGISLGNSEADQLLEAARA 502
 QY 328 -DTE-----ICT-----KDGE-----TPLIKATKRNIEVEVLELLDCAKAVSAVDKGDTP 372
 DB 503 GDEYEVKKTCTVQSVNCRDLEGROSTPLHFAAGYNVSVVEYELLQHGADVHAKDKGGLVP 562
 QY 373 LHTAIRGRSKLAELLRN 391
 DB 563 LHNACSYGHVEYVALLVKH 581

RESULT 12
 RNA_HUMAN STANDARD; PRT; 741 AA.
 AC 005823;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE 2-5A-dependent ribonuclease (EC 3.1.26.-) (2-5A-dependent RNase)
 DE (ribonuclease L) (RNase L) (ribonuclease 4).
 GN RNASEL OR RN54.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN 11
 RP SEQUENCE FROM N.A., AND MUTAGENESIS OF LYS-240 AND LYS-274.
 RC TISSUE=Kidney;
 RX MEDLINE=93201598; PubMed=7680958;
 RA Zhou A., Hassel B.A., Silverman R.H.;
 RT "Expression cloning of 2-5A-dependent RNase: a uniquely regulated
 RT mediator of interferon action.";
 RL Cell 72:753-765(1993).
 RN 12
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=20515260; PubMed=11063255;
 RA Zhou A., Nie H., Silverman R.H.;
 RT "Analysis and origins of the human and mouse RNase L genes: mediators
 RT of interferon action.";
 RL Mamm. Genome 11:989-992(2000).
 RN 13
 RP CHARACTERIZATION OF RNASEL ACTIVITY.
 RX MEDLINE=94245737; PubMed=7514601;
 RA Dong B., Xu L., Zhou A., Hassel B.A., Lee X., Torrence P.F.,
 RA Silverman R.H.;
 RT "Intrinsic molecular activities of the interferon-induced 2-5A-
 RT dependent RNase.";
 RL J. Biol. Chem. 269:14153-14158(1994).
 RN 14
 RP FUNCTION, AND SUBCELLULAR LOCATION.
 RX PubMed=11585831;
 RA Salehzadeh T.;
 RT "The 2-5A/RNase L/RNase L inhibitor (RNI) pathway regulates
 RT mitochondrial mRNAs stability in interferon alpha-treated H9 cells.";
 RL J. Biol. Chem. 276:48473-48482(2001).
 RN 15
 RP ERRATUM.
 RA Le Roy F., Bisbal C., Silhol M., Martinand C., Lebleu B.,
 RA Salehzadeh T.;
 RL J. Biol. Chem. 277:13354-13354(2002).
 RN 16

RP REVIEW.
 RX PubMed=9856285;
 RA Castelli J., Wood K.A., Youle R.J.;
 RT "The 2-5A system in viral infection and apoptosis.";
 RL Biomed. Pharmacother. 52:386-390(1998).
 RN 17
 RP MUTAGENESIS OF LYS-392.
 RX PubMed=9862963;
 RA Dong B., Silverman R.H.;
 RT "Alternative function of a protein kinase homology domain in 2',
 RT 5'-oligoadenylate dependent RNase L.";
 RL Nucleic Acids Res. 27:439-445(1999).
 RN 18
 RP MUTAGENESIS OF HIS-583; PRO-584; TRP-632; ASP-661; ARG-667 AND
 RP HIS-672.
 RX PubMed=11333017;
 RA Dong B., Niwa M., Walter P., Silverman R.H.;
 RT "Basis for regulated RNA cleavage by functional analysis of RNase L
 RT and Irep.";
 RL RNA 7:361-373(2001).
 RN 19
 RP VARIANTS SER-59; PHE-406; GLN-462 AND GLU-541.
 RX PubMed=11941539;
 RA Roekman A., Ikonen T., Seppaelae E.H., Nupponen N., Autio V.,
 RA Mononen N., Bailey-Wilson J., Trent J., Carpen J., Matikainen M.P.,
 RA Koivisto P.A., Tammela T.L.J., Kallioniemi O.-P., Schleutker J.;
 RT "Germline alterations of the RNASEL gene, a candidate Hpc1 gene at
 RT 1q25, in patients and families with prostate cancer.";
 RL Am. J. Hum. Genet. 70:1299-1304(2002).
 CC 1- FUNCTION: Endoribonuclease, mediator of interferon action, which
 CC play a role in mediating resistance to virus infection and
 CC apoptosis. Might play a central role in the regulation of mRNA
 CC turnover.
 CC 1- CATALYTIC ACTIVITY: Cleaves 3' of UpNP dimers, with preference for
 CC UU and UA sequences, to sets of discrete products ranging from
 CC between 4 and 22 nucleotides in length.
 CC 1- COFACTOR: Optimal RNA cleavage rates requires the presence of
 CC either manganese or magnesium and ATP.
 CC 1- ENZYME REGULATION: After binding to 2-5A (5'-phosphorylated 2',5'-
 CC linked oligoadenylates) the homodimerization and subsequent
 CC activation occurs. Inhibited by RNase L inhibitor.
 CC 1- SUBUNIT: Monomer (inactive form) or homodimer.
 CC 1- SUBCELLULAR LOCATION: Cytoplasmic and mitochondrial.
 CC 1- TISSUE SPECIFICITY: Highly expressed in spleen and thymus followed
 CC by prostate, testis, uterus, small intestine, colon and peripheral
 CC blood leukocytes.
 CC 1- INDUCTION: By interferons.
 CC 1- DOMAIN: The kinase domain allows the homodimerization.
 CC 1- DOMAIN: The nine ankyrin repeats also called 2-5A sensor
 CC constitute the 2-5A binding domain.
 CC 1- DISEASE: Germline defects in RNASEL are linked to cancer-
 CC susceptibility for HPC1.
 CC 1- SIMILARITY: CONTAINS 9 ANK REPEATS.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: L10381; AAA18032.1;
 DR HSSP: P80144; 2MTO.
 DR Genew: HGNC:10050; RNASEL.
 DR MIM: 180435;
 DR InterPro: IPR002110; ANK.
 DR InterPro: IPR000719; Euk_pkinase.
 DR Pfam: PF00063; ank; 8.
 DR ProDom: PD000001; Euk_pkinase; 1.
 DR SMART: SM00248; ANK; 8.
 DR PROSITE: PS0088; ANK_REPEAT; 6.

Tue Jul 1 15:13:30 2003

us-10-021-571-4.rsp

Page 19

Db 291 LLEKGA-----DKCDDLWMIARRNHDYHLVKLLPYVANP 326

Search completed: July 1, 2003, 14:35:24
Job time : 38 secs

Query Match	6.2%;	Score 552.5;	DB 2;	Length 1856;
Best Local Similarity	19.6%;	Pred. No. 1.2e-22;		
Matches 373;	Conservative 287;	Mismatches 659;	Indels 581;	Gaps 72

QY	519	FVHPMLGIAVSLPLALLIFVLYYVGGREDESMNMAVMLSTRARHIGLELLK	577
Db	657	KQANGMLGKKSGLTPLHLV-----AQEG-----HVPADVLLIH	690
QY	578	-----LMFNPPELPEQTTKALPYEFLF-----TDYNRLSSVG-----GETSIA	616
Db	691	GWVAVATRMGVTPPLHVASHGNKIKVYFLLQHOADVNAATKTKIYSPHQAQOQHTDVL	750
QY	617	EMI-----ATLSDACEREGELA-TLFFVFKTEDTQGGKKKKTKCCLPSV	662
Db	751	TLLKNGASPNVSSDGTTPLAIAKRLGIVISVTVLKV-TDET-----SEV	796
QY	663	IFELFICLISITLLAIFRPDRPHL-----TVNAVLSIASVVG---LAFVINCWTWQ	714
Db	797	L-----VSDHNRMSFPEYDEIL-DVSEDEGBELISFAERR-----	832
QY	715	VLDLSLNSQRRKLHNAASKLHLKLSSEGMVKVLCVEVLAIRMAKTIDTSFOTMOTRLVIT	774
Db	833	-----DSRVDBEKLLDFVKIDQYVESAIRIPIC-----AMEYIVINSEBEGASKEY	884
QY	775	DGLDACEODKVIQMDIVYVLFSS-KGFPLAIFASDP-----HIITK-----	814
Db	885	DEDSLIPSSPATETSDNISPVASPVHGFELFSEFVADARGSGMSGRHNGILRVIPRTCA	944
QY	815	-----AINONNSYLRDSNINGHDYMRN-TVHLPYEL-NSRGIS	851
Db	945	APRITICRLYKQKISTPPPLAEEGBGLASRIIALGPGAOPLSVYVEIPIHFSHORG-	1007
QY	852	NAKRELVTASATNDVPCSDTTGIOEDADRVYONSLEMTKIGSKTALNRRDTYRRRQMO	911
Db	1003	-----DRELVYLSERSEGV-----WKEHRSRGYESIYDQIILN-CMDELSGSLELEKKRVC	1052
QY	912	RITTRMSDITLKLVT---EDWESDISPQIMRNLINIVSYTRLLRANQISFMDRLAS	968
Db	1053	RITT---TDEPLVFMISRLCOD-VDIIGPEG-----GSLKSLVPLVQATFEPENAVTK	1102
QY	969	WINTLEQWYFRSMILYLEETEGIPPOOMTLK-----TYERISK-----	1008
Db	1103	KYKLAQ-----AQYVPDELYTKLLGNQATSPITYVYPRRRKFRHPIGLR	1146
QY	1009	NIPPTKQVEPLLEIDGINFEVYF-----LSRTPVLVADVKYFLPCTV	1053
Db	1149	IPLPSSWTDNPRDSEGDJTTSLRLICSVIGGTDQAWMEDJTGTKLKYVANECANF---TT	1205
QY	1054	NIDPKLRELIADYRARARQIISIGLAIPPLPLHBGPPRAPSGISQPPSCSSTFRNGPPA	1113
Db	1206	NV-SARFWLSCDPTAAVNAVFALLKYEL-----TAPYMAKEV	1243
QY	1114	GGVNSPQPHSSYYSGMTGQHPFNRRSGPAPGV---VLLNSLVNDAVCE-----K	1163
Db	1244	-----IFAKANDPREGRLCRYCTMDTKDVKVTKTEJQHBENFVEAR	1281
QY	1163	LKQIEGLDOSMLPOYCTTIKKNANINGRYLAOCNIDELKKEMMNFGDWHLFRSTYLEMRN	1222
Db	1282	SRDIELEGMSL-----FAELSG-----NLVYPKKAQQRSPHFQSPFENRLAM-	1322
QY	1223	ASHVYVDEPRRLSSSSSGPARPHEPARARASHNE---LPHTELS-----	1263
Db	1326	-----PVKVRDSSREPGSSLSFLRKAKKEDYTOHILCHNITIMPCCANGSGAEDRR	1376
QY	1264	-SOTPYTLNFSFEELNLTGLDEGAPR-----HSNLSMOSQTRRRP-SLSSLSNSD	1311
Db	1377	KRPPTPLALRYLSLSTSTPGSLSTGEQAOEMKAAVYSEHLGISMWELARELOFVEDIN--	1433
QY	1312	SSIEISKLTDKVQAVRYADRYEYIAQMSOL---EGQPGSTTISGRSSPHSTYYMGQSSS	1367
Db	1434	-----RIRVENPMSILLEQSVALLNTLVIREG-----QANMENNYTLAQOSID	1475
QY	1368	GGSHSNLEQEKGDSEPKPDGKSKFLMKRGVDIYSSSGVSTNNASPLDITTEDE--	1422
Db	1476	RGEIVNMLEGSGRQSRMLKPD-----RRHTDRDYSLSPSQMNGYSSL---ODETL	1522

QY 1426 -----KSDGSSKLT-----LPKSSSRSLFQDMLKSGLRKYOKLPDE 1467
Db 1523 SPASIGCALSSPLRQDQWNEVAILDAIPLAITEHDITLWMSDMQWASAGLTPLSLVRAED 1582
QY 1468 DE---SGTEESDNTPLDKDRKAQGERVYKSPESHAPRTIFIKAYEYLDALDK 1524
Db 1583 SLSECKSAEDSDAT-----GHEWKEGALSEPRGPE-----LQSLTEVEDDVTDS 1628
QY 1525 KDS-----DSGVRSESPSPNLSINEVADDSQLEKANLIELEDSSHSGKGIPISTL 1576
Db 1629 DATNLDLLEQEEQSEEEKLPQSKRODDATGACQ-DSENEVSLVSGHQRQARITHS 1686
QY 1577 SGLDPIIARMSICSEDK-----KSPSECSLIASPEENW 1611
Db 1687 -----PIVSOYTERSDRLQDMDADGSLVYLQDAQGSW 1721

RESULT 6
A35049
Ankyrin 1, erythrocyte splice form 2 - human
N.Alternate names: ankyrin 2.1, erythrocyte; ankyrin-R
C.Contains: ankyrin 2.2, erythrocyte
Species: Homo sapiens (man)
C.Date: 27-Jul-1990 #sequence_revision 01-Oct-1992 #text_change 04-Sep-1998
C.Accession: A35049
R.Lamert, S.; Yu, H.; Prechal, J.T.; Lawler, J.; Ruff, P.; Speicher, D.; Cheung, M.C.; K
Proc. Natl. Acad. Sci. U.S.A. 87, 1730-1734, 1990
A.Title: cDNA sequence for human erythrocyte ankyrin.
A.Reference number: A35049; M01D:90175370; PMID:1689849
A.Accession: A35049
A.Status: preliminary
A.Molecule type: mRNA
A.Residues: 1-1880 <LAMP>
A.Cross-references: GB:M28880
C.Genetics:
A.Gene: GDB:ANK1; ANK
A.Cross-references: GDB:118737; OMIM:182900
A.Map position: 8p11.2-8p11.2
C.Superfamily: ankyrin repeat homology
C.Keywords: alternative splicing; cytoskeleton
F:2-1880/Product: ankyrin 1, erythrocyte form 2 #status predicted <MAT>
F:2-1513,1676-1880/Product: ankyrin 2.2, erythrocyte #status predicted <MA2>
F:44-76/Domain: ankyrin repeat homology <AN01>
F:77-109/Domain: ankyrin repeat homology <AN02>
F:110-142/Domain: ankyrin repeat homology <AN03>
F:143-171/Domain: ankyrin repeat homology <AN04>
F:172-204/Domain: ankyrin repeat homology <AN05>
F:205-237/Domain: ankyrin repeat homology <AN06>
F:238-270/Domain: ankyrin repeat homology <AN07>
F:271-303/Domain: ankyrin repeat homology <AN08>
F:304-336/Domain: ankyrin repeat homology <AN09>
F:337-369/Domain: ankyrin repeat homology <AN10>
F:370-402/Domain: ankyrin repeat homology <AN11>
F:403-435/Domain: ankyrin repeat homology <AN12>
F:436-468/Domain: ankyrin repeat homology <AN13>
F:469-501/Domain: ankyrin repeat homology <AN14>
F:502-534/Domain: ankyrin repeat homology <AN15>
F:535-567/Domain: ankyrin repeat homology <AN16>
F:568-600/Domain: ankyrin repeat homology <AN17>
F:601-633/Domain: ankyrin repeat homology <AN18>
F:634-666/Domain: ankyrin repeat homology <AN19>
F:667-699/Domain: ankyrin repeat homology <AN20>
F:700-732/Domain: ankyrin repeat homology <AN21>
F:733-765/Domain: ankyrin repeat homology <AN22>
F:766-798/Domain: ankyrin repeat homology <AN23>

Query Match 6.2%; Score 552.5; DB 2; Length 1880;
Best Local Similarity 19.6%; Pred. No. 1.2e-22;
Matches 373; Conservative 287; Mismatches 659; Indels 581; Gaps 72;

QY 21 LKALLEKCDVDERBECGQTPIMIAEOGNLEIVKELKNGANCLEDDLMWTALISASK 80
Db 94 VRELNVYGANVNAOSOKGFTPLYMAOEHNLEVVAFLELNGANOVNATEDGFTPLVALQ 153

QY 81 EGHVHIVEELKCGV-----NLEHRDGMGTALMA 111
Db 154 OGHEVNVVAILINTGYKQVRLPALHIAARNDTRIAVLLONDPNVDLSKGTFTPLHIA 213
QY 112 CYKGRDVEVELLSHGANSVT-----GLQY-----SV 139
Db 214 AHENENLVAGLLINRSGSVFTFQNGITPLHIASRGVNIYVRLLDRAQJLETKYKDEL 273
QY 140 YPTIWAAGRHADIYVHLLONGAKVNCSDKYGTTPLVMAARKHLECYKHLIAMGADV- 198
Db 274 TPLHCARNGHVRISELIDHGAPIQAKTKNGLSPIHMAOGHLDCCVLLLODYADEID 333
QY 199 -----OEGA-----NSMTALIVAKGQTSQVEIKLR 226
Db 334 ITLDHLTPLVAAACHHRAKVALLDKGAKPNSBALNGFTPLHIAKKNHVRMELLT 393
QY 227 NPVNLITDKDNTALMASKEGTEIVODLLDAGTYVNIIPDRSGDTVLIGAVRGHVEIV 286
Db 394 GASIDAVTESGLPLVHVASPMGHLPIYKNLQGASPNVSNVETPLHMAARAHTVEVA 453
QY 287 RALLQKADIDINGODNKTALVYAVEKGNATVRODLOCPTEICTKGEFTPLKATM 346
Db 454 KYLLONKAKVNAKAKDQPLHCAARIGHNMVKLLIENNANPLATFGHPLHIAARE 513
QY 347 RNEIYVELLDKGAAYAVDKGDTPLHIAIRGRSKLAEILLRNPKDRLILYRPNKAGE 406
Db 514 GHVETVALLLEKERSQACMKTKGFTPLHVAAKGKRVAVELLE--RDA-----HRAAK 567
QY 407 ---TPYIDCSHOKSLITQIFGARHLSPEETDGMIGY-DLYSSALADILSEPTMOPIC 462
Db 568 NGTLPLHVAVHNHNLIDVKLLLRPGSGPSHPANN--GYPLHIAAKON-----QVEYA 618
QY 463 VGLYAGWCSGSKPLTKLEDEMKTPAGOOLEPFGOSMLIVFTLLCGS---LGILFA 518
Db 619 RSLQYGGSANASV-----QGYPLH-----LAAQSHAMVALLS 656
QY 519 FTVHPNLGIIVSLFALTYFETVYFGRRGESMNNAMVLSRLAHIGYELLRL 577
Db 657 KQANGNLGNKSGTLPLHV-----AQGS-----HVPAVDLIRK 690
QY 578 -----LAFVNPPELPEQTTKALPYRPLF---TYNRLSSVG-----GETSLA 616
Db 691 GVMVDAITRMGYTPLVASHYGNIKLVKFLQDAQVNAKTKGYSPLHQAQOGHTDIV 750
QY 617 EMI-----ATLSDACERESEFLA--TRLPRVETEDTQCKKKKTKCCPSFV 662
Db 751 TLLKNGASPNVSSDGTTPLAIAKRLGITSVDYLVKV--TDET-----SFV 796
QY 663 IFLEITGCIISGTLIAIFVNDPKHL-----TVNAVLIASVYG---LAFVLCRTWMO 714
Db 797 L-----VSDKHRSFPETVEIL--DVSEDEGBELISFKAERR--- 832
QY 715 VLDSLNSQKRRLHNAASKRLKLSSEFMVNLCEVEIARMAKTKIDSTFQNTRLVYII 774
Db 833 ---DSRDVDEKEILDFVFKLDQVESPALPRIC-----AMDETIVINSEDEQASKEY 884
QY 775 DGLDACRQDVLQMLDTRVLFs--KSPFAIASPDP-----HIITK----- 814
Db 885 DEBSLIPSSPATEISDNISPVASPVHTGFLVSPVNDARGSMGSRHNGRLAVIIPRTCA 944
QY 815 -----AINQNLNSYLKDSNINSHDYMKN--IVHLVPL--NSRGIS 851
Db 945 APTRTCLVVKPQKLTSPPLAEEGLASRIIALGPTGAQFLSPVIVEIPHFASHRG-- 1002
QY 852 NAKKFLVTSATNGDVPSCSDTGTIOEDADRVSQNSLCEBMTKLSKALNRDPIRRRQMO 911
Db 1003 -DRELVLVRSSENSV-----WKHRSTYGSYSDQILN--GMDEGLSLELEKRCVC 1052
QY 912 RTTRQMSFDUTKLAVT---EDWFSDISPQTMRLINIVSYGRLRANQISFPMRLAS 968
Db 1053 RIIT--TDFFLYIVIMRLCOD--YDIIGPEG-----GSLKSLVPLVQATPPEANAYIK 1102

QY 969 WINTEQMPYRTSWLLILEETGIPDOMTLK-----TIVERISK----- 1008
 Db 1103 RKVLALQ-----AQPVPDELVTKLIGNQATFSPITYVEBRARRKFRPIGLR 1148
 QY 1009 -NIPPTKVEPPELEIDGIRNFEYV-----ISSRPVLVARDVYKVLPCYV 1053
 Db 1149 IPLPSPWTDNRDGEEDTSLRLICSVIGGTDQOMEDITGTTKLAVANECANF---TT 1205
 QY 1054 NLDPKLEIIVADVRARPOISIGGLAVPPLHGGPPRAPSGYQSPVCSSTFNQPEFA 1113
 Db 1206 NV--SARFWDSCRTAEAVAFATLLYKEL-----TAVPYMAKFFV 1243
 QY 1114 GGVSVPQPHSSSYSGMTGPQHPFYNRKSGPAPGV--VLNLSLVDAVCE-----K 1162
 Db 1244 -----IPAKMNDPREGRRCYCMDDKVDKTLTQGHENFEVAVR 1281
 QY 1163 LKQIEGLDSMLPOYCTTIKANKINGRYLAQCINDELKKEKNNMFGWHLFRSTVLEMRN 1222
 Db 1282 SRDLEVELEGMSL-----FAELISG-----NLVYPKKAQORSFHFOSFRENRLAM-- 1325
 QY 1223 AESHVVEDPRLFLESSSGPAPHGEPAPRASHNE-----LPHTELS----- 1263
 Db 1326 -----PVKVDSSREPGSGSLFRLKANKYEDTQHLCHLNTITWPPCAKSGGADRR 1376
 QY 1264 -SQPTIINFSEELNTLGLDEGAPR-----HSNLSWQSOTRTP-SLSLSNSOD 1311
 Db 1377 RTPPLALRYSLISESTPGSLSGTEQAEKMAVISEHLGLSMELARELOFSVEDIN--- 1433
 QY 1312 SSIISKLTDRVQAEYRAYREYIAOMSOL-----EGGPGSTISGRSPSTHYMGSSS 1367
 Db 1434 -----RIRERNNSLLEQSVALLNLMTVIREG-----QANMELTYALDSTD 1475
 QY 1368 GGSHTSNLEDEKGDSEPKPDGKRKSPLMKRGVDYIDYSSGSVSTNDASPLDPTTEDE-- 1425
 Db 1476 RGEIVNMLEGSGRQSRMLKPD-----RRHTROYSLSPSQMNQVSSL-----QDEL 1522
 QY 1426 -----KSDQSGSL-----LPKKSSEBSLSQTDILKLGSLRQKLPSPDE 1467
 Db 1523 SPASLGCALSSPLEADYDWMVEVALIDAIPLAETHDMLKMSDQVMSAGLTPSLVLAED 1582
 QY 1468 DE--SGTESDNPPLKDKDKRAKGKVERVPRKSPHSABPIRTFLKAKELYSDALDK 1524
 Db 1583 SSLECSKAEDSDAT-----GHEWKLVEALSEERGPPE-----LGSLELVEDDTVDS 1628
 QY 1525 KDSS-----DSGVSSSSSPNHSJLANEVADDSQLEKANLIELEDSSHGKRGIPHS 1576
 Db 1629 DATMGLIDLLEQEGGRSEKELPQSKRQDDATGAGQ--DSENEVSLVSGHQRGARITHS- 1686
 QY 1577 SGLDPIIARMSICEK-----KSPSECSLIASSPEENW 1611
 Db 1687 -----PTVSOVTERSQDRLODMWDADGSIVSLDAAOGSW 1721

RESULT 7

AE2149
 hypothetical protein all2748 [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp.
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
 C:Accession: AE2149
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuriitz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
 DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena Res. 8, 205-213, 2001
 A:Reference number: AB1807; MUID:21595285; PMID:11759840
 A:Accession: AE2149
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-426 <KOR>
 A:Cross-references: GB:BA000019; PIDN:BAW74447.1; PID:g17131841; GSPDB:GNO0179
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: all2748

Query Match 6.1%; Score 544; DB 2; Length 426;
 Best Local Similarity 37.2%; Pred. No. 3, 2e-23;
 Matches 149; Conservative 64; Mismatches 180; Indels 8; Gaps 6;
 QY 9 VINVEEENIPALKALECKDYDERNEGQFPLMTAAAGNLEIYELIKKANGCNL-E 67
 Db 8 LKVAKSGDIKIGALLAGVGVDDICDRGTTALMRAINDTEIYRSLLDGGANVNLAR 67
 QY 68 DLDMNTALISAKREGHVHYEELLKGVNLEHRDMGWTALMACYKRGTDVVELLSHG- 127
 Db 68 KRYGTLALMLAASANOYDIQVLLISGAAVNATNEDGRTALMAALAKGVAVARVLLAG 127
 QY 128 ANPSVGLQSVYPIITMAARGHADYVHLLDNGAVNCSDKYGTPLYMAARKHLECY 187
 Db 128 ADVNTDKDDT-ALKLAVKRGGAAYVOLLPSGADANCEDEGETLLMLADSGHGVV 186
 QY 188 KHLAMGADVDEGASMTALIYAVKGYTQSVKEILKRNPNVNLTDKDGNTALMASKE 247
 Db 187 QVLLAGVDPVNEQNDGGTALLAAVAAGNCAIAKILLDGGADVNRHDDGSESALHAYE 246
 QY 248 GHTEIYODLIDAGTYVNIIDPRSGDVLIGAVRGHVEIYRALLQYADIDIRGQD-NKTA 306
 Db 247 GYDVYQVLLNQGANTQINKKLGDTPLVLAALQGHQIYETILL-KYG-ANVHGDMIGETP 304
 QY 307 LYMAVERGNATVVRDILQCNPDTEICTKGETPLIKATMBRNEVEYELLDDGAVYSND 366
 Db 305 LTILAAGGHATYRIILDDYGANANIPASOGKTLIATERNHPGVQLLLAGANVYOD 364
 QY 367 KKGDPLHAIIRGSRKLAELLRNPKDRLLYPRPKAGET 407
 Db 365 SVGATALLMAASGYNKVVQJLLEGADTNL--KNRGYTT 402

RESULT 8

S37431
 ankyrin 2, neuronal long splice form - human

N:Alternate names: ankyrin B, 440K splice form; ankyrin-B; brain ankyrin; non-erythro
 N:Contents: ankyrin 2, short form
 C:Species: Homo sapiens (man)
 C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 13-Aug-1999
 C:Accession: S37431; A39643; B39643; A40334; A49462; S14569
 R:Chan, W.
 submitted to the EMBL Data Library, September 1993
 A:Reference number: S37431
 A:Accession: S37431
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-3924 <CHA>
 A:Cross-references: EMBL:Z26634; NID:q406287; PIDN:CAA81387.1; PID:q406288
 R:Otto, E.; Kunimoto, M.; McLaughlin, T.; Bennett, V.
 J. Cell Biol. 114, 241-253, 1991
 A:Title: Isolation and characterization of cDNAs encoding human brain ankyrins reveal
 A:Reference number: A39643; MUID:91302466; PMID:1830053
 A:Accession: A39643
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-2077 <ORT>
 A:Cross-references: GB:X56957
 A:Accession: B39643
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1443, 3585-3924 <ORT>
 A:Cross-references: EMBL:X56958
 R:Use, W.T.; Menninger, J.C.; Yang-Feng, T.L.; Francke, U.; Sahr, K.E.; Lux, S.E.; Wa
 Genomics 10, 858-866, 1991
 A:Title: Isolation and chromosomal localization of a novel nonerythroid ankyrin gene.
 A:Reference number: A40334; MUID:92009921; PMID:1833308
 A:Accession: A40334
 A:Molecule type: DNA
 A:Residues: 463-474, 'PE', 477-495 <TSE>
 A:Cross-references: GB:M37123; NID:9178647; PIDN:AAA62828.1; PID:g178648
 R:Chan, W.; Kordeli, E.; Bennett, V.


```

QY 876 ---EDADRVSQNSIGEMTKGSKTALNRDPT---RRQOMORTTROMSFDLTL---V 927
Db 1184 PVPETVAKI-----LGNKATFSPITVPEBRKRKRKPT---MTIPVPPSGEGV 1231
QY 928 TEDWSDISDPTMRRLNIVSGTGLRANQISFWMDLAAMINLTCOWP-----Y 978
Db 1232 SNGYGDATP-NLRILCSITGISTP-----AQMEDITGTPPLTFIKDCVSF 1276
QY 979 RTS-----WL-----ILYEETEGIPDOMTKTI-YERISKNIPTTKDVEPLIEDIDIR 1027
Db 1277 TTNSARFRLADCHQVL---ETVGLASQLYRELICVPMKAFVFAKNDP---VESLIR 1330
QY 1028 -----NFEVFLSGRPPVLARQVKKVLPCTVMDPLR---ELIDVR 1067
Db 1331 CFQMTDDRVKTLQOENFEVYARSK-DIEVLEGKPIVDCYGNAPLTKGGQDVENFY 1389
QY 1068 AAREQISIGLAPPLPLHEGPPRAPSGY-----SQPPVCSSTSFNGPFAGVVSQPH 1122
Db 1390 SFKEN-----RUPFSIKIRDSQEP---CGRLSF----- 1415
QY 1123 SSIYSGMTPOHPFYNRSGAPGPVYLLNSLNDVANCEKIQLEGLDQSMLEPOYCTTIK 1182
Db 1416 -----LKEPKTKTKGLPQTAV----- 1430
QY 1183 KANINGRYLAQCNIDELKEMNMNGDHLFRSTVLEMRNASHVVPEDPFLSESSGP 1242
Db 1431 -CNLNTITPAHKKAADRRS-----FASLALRRKYS-----YLTPSKSP 1471
QY 1243 APHGEPANR-----ASHNELPTELSQPTPTLNFSEFELNTLGLDEGAPRHSNIS- 1293
Db 1472 Q---SPCERTDIRMAIVADHGLSWTELARE---LNFVSDEINOIRVEN---PNSLSIQ 1521
QY 1294 -----W---QOFRRTPSLSLNSODSIEISKLTKVQAEY-----RDATREYIAQ 1337
Db 1522 SMMLKKNVTRDGNKATDALTSLYTKINRDIYTLLECPITFDIGNISGTSFADENNVF 1581
QY 1338 MSQLEGGFG-----STTISGRSSPHSTYYMGSSSGSITHS--N 1374
Db 1582 HDPVGDHPSFQVELETMPGLYWTAPNPFOODHFDISSISPRFRLSDGLVPSQGN 1641
QY 1375 LEQENG-----KSS---EKPRDGKRSFLMKRGDVIDYSSSGVSTNDASPLDPTIED 1424
Db 1642 IEHPTGGPPVYTAEDTSLSDSKMD----SVTVTDPPADPLD--VDES 1682
QY 1425 EKSDSGSKL-----LPKKSSEKRS--LFOTDLKLGSLRGYOKLPESDDESGTE--- 1473
Db 1683 QUKDLQSGEACQACMAVSFGIRPDGQAPLAPQTKKVMSSQOEKKGSGDEVEYTEKV 1742
QY 1474 -----ESDNTPLKDDKRAEGVERVP-----KSPHSAAE-PLRTFIK 1512
Db 1743 KSLFEDIQLEVEAEEMETDOGAAMLNRVQRAELAMSLAGMONTPEGSLESAPAQ--- 1799
QY 1513 AKETYSDDLKKKSSDSGVRSS-----ESSPNHSLNHEVADSDQLEKANLIE 1560
Db 1800 -RRLTGLDLRLDSSDQADRSITSYLTGEPKILEANGNHTA--EVIPKA--KAKYF 1852
QY 1561 LEDDSHSGRGIT-----PHSLGLDPIIARMSICSEDKKSPSECS--FIASSPEMN 1611
Db 1853 PESQNDIGKOSIKELKAKTKHCCGTEBPV---SPLTAYOKSLEETSKLVIEDAPKCV 1908
QY 1612 PACQAKVAMLRTPSTVTVLNNNSAPANRANQNFDEMEGIRETSQVILRPSSPNPTTONE 1671
Db 1909 PVGMK--KMTRTTAD-----GKARLNIQEBEG-----STRSEBK--QGE 1943
QY 1672 NIKSMTHK 1679
Db 1944 GYKVKTKK 1951

```

RESULT 10
T42715
ankyrin 3, splice form 3 - mouse
C:Species: Mus musculus (house mouse)

```

C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000
C/Accession: T42715
J:Peterson, L.T.; John, K.M.; Lu, F.M.; Elcher, E.M.; Higgins, A.; Yialamas, M.; Turtzo
J. Cell Biol. 130, 313-330, 1995
A:Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin ge
the repeat domain.
A:Reference number: 222237; PMID:95340633; PMID:7615634
A:Accession: T42715
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1940 <PEP>
A:Cross-references: EMBL:U40632; NID:9710548; PID:9710549; PIDN:AAB01604.1
A:Experimental source: strain C57BL/6J; kidney
C:Genetics:
A:Gene: Ank3
A:Map position: 10
A:Introns: 834/1
A:Superfamily: ankyrin; ankyrin repeat homology
C:Keywords: alternative splicing

Query Match 6.08; Score 529.5; DB 2; Length 1940;
Best Local Similarity 20.18; Pred. No. 2,6e-21;
Matches 431; Conservative 280; Mismatches 673; Indels 763; Gaps 90;

QY 14 EENIPALKALKECKDVERNECGOTPTMIAEQNLTEVELKINGANCNLEIDNMT 73
Db 66 KEGHEVEVSELLDREANVDAATKKGNTALHISLAGQAEVYKLVYNGANVNAQSGNGT 125
QY 74 ALISAKKEGHVHIVEILKCGVNLHHRDMGWTALMACYKGRDVEYELL----- 124
Db 126 PLYMAQENHLEVRFLDNGASQSLATEDGFTPLAVALQGDYVSLLENDTKGVR 185
QY 125 -----SHGANSVYGL----- 135
Db 186 LPLALHAARKDPTKAAALLQNDTNADVESKSGFTPLHAHNGINVAFTLLNRAAD 245
QY 136 ---QYSYPLIIMAGRGHADIVHLLONGAKYVCSKYGTPPLVMAARKG----- 182
Db 246 FTARNITPLHVASKRGKMANVKKLLDRGAKIDAKTRDGLPLHCGARGSHQVVEMLD 305
QY 183 -----HLECYVHLLAMGADVQEGANSFTALIVAKGYTOS 219
Db 306 RSAPILSKTKNGLSPLMATQGDHLNLCYDILLQHNVPVDVYNDYLTALHVAHCGHYKV 365
QY 220 VKELIKNPVNLTDGNTALMIASKE----- 247
Db 366 AKVILLKKSAPNAKALNGFTPLHIAKKNRIRYEMELLKHGASIQAVTESGLPTIHAUF 425
QY 248 -GHEIVQDLDLAGTYVNIIPDRSGDTVLIGAVGHEVETVIRALLQRYADIDIRGDNKTA 306
Db 426 MGHVNIYSQLMHGGASPNFTNVGFTALHMAARSQAEEVYRVLYDQGAQVEAKAKDQTP 485
QY 307 LYMAVEKGNATMRDILQCPDTEICTKGEFTPLIATKRNIEVEVELLDGAVYSAVD 366
Db 486 LHISARLGRKADIVQOLLQOGASPNAAFTSGYTPPLHAAREGHEDVAALFDIGASLSTYT 545
QY 367 KKDPTPLHAIIRGRSKLAELLR-----NPKDRLL----- 398
Db 546 KKGFTPLHVAAKGKLEVASILLQKSASPDAGKSGLTPLHVAAHNDQKVALLLLDGA 605
QY 399 -YRPNAKGETPYNIDC-SHOKSILTOI--FGAR-----HSPETEGDMIGY 441
Db 606 SPHAAAKNGYTPHIAAKKQMDIATSLLEYGADANAVTROIASVHLAAQGHVDMVSL 665
QY 442 DLYSALADLSEPTQIP-----PICVGLYQW 469
Db 666 LLSRMANVNLNSKGLTPLHLAAQEDRVNAVVEVLVNOGAHVDAQIKMGYTPPLHVGCH--Y 723
QY 470 GSGK--SFLK---KLEDEKKT-----FAGQOIEPLFOFWSWLVLTLLLCG---LG 514
Db 724 GNIKIYVFLQHSAAKNAKKNKYTALHQAQO-----GHTHILN 763
QY 515 LLEFAIVHPN-----LGIAVSLFIALLYIFIV----- 543

```

```

Db 764 VLLQNNASPELTVNGNTALAIARLGIYSVDTLKVTEIMTTTTEKHNVPETM 823
QY 544 --LYFGGRREGSEMNAMVLTSLARHIGYLELLKL-----MFVNPRLPQTK 592
Db 824 NEVLDMDDGDGCTCFKIPKVO-----EVLVASEDAITDDTKYGLPDLKELGDD 875
QY 593 ALP-----VRELFED-----YNRLSVGGETSLAEMIALSDACERFGL 633
Db 876 SLPAEGYVGSGLARASLSLFSDDRSYTLNRSYARDSMMEILLVP-----SKQHIL 930
QY 634 ATPLFRFKTEDYQGGKKKKKCTCLPSVYLFELICGIIIGT--TLALFRVDPKHLTVN 691
Db 931 FTREF--DSDILRHVMAADT-----LDNVLVSSPVHSGFLVSPMDARCGSKR 978
QY 692 AVLIASVYGLAFVL--NCRPMOVLDLSNOKRRLHN-----AASKLKLS 739
QY 979 G-----SRHGMGLIIPPKCTAPRTICLV--KHKLANPPMYEGEGLASRLVEMGP 1031
Db 740 EG-----FMKYLKCEVELMARMAKTIDFTONQ-----TRLVIIIGLAGE 781
QY 1032 AGAQLGPIVVEIPIPHGSMRGKRELLVLRSENGETWKEHQFDCKMEDLALLNGMD--E 1089
QY 782 QDAVLQMDTVRY--LFSKGFPAIFASDPHIIITKAINQNLNVLDNSNGHDYKNTV 839
Db 1090 ELDSPELGTGRICRIITK--DEPOYPA-----VSRIKOESNOIGEGGI----- 1133
QY 840 HLEVFINSRLSNARKFLV--TSATNGDVPSCDTGTGQ-----EDADRVYSONSLGEMTK 893
Db 1134 -----LSTYTPVLVQASPEBALIKRIIRVGLQADVPREYTKI-----L 1173
QY 894 GSTALNRRDTY--RRROMRTTROMSFDLKL--VTEWFSDISQPMRLNIVS 948
Db 1174 GNAKATSPYIVTVEPRRRKFKPIT--MTIPVPPSGEVSNGYKGDATP--NLRLCSTIG 1230
QY 949 VTRGLRANQISFNMRLASWMLTROWP-----YRTS-----WL-----IYLEE 989
Db 1231 GTPP-----AOWEDITGTPPTFTIKDCVSTFTNVASARFWLADCHQVL--E 1273
QY 990 TEGIPDMTLKTI--YERISKNIPTTKDVEPLEIDGDIR-----NFEV 1031
Db 1274 TVGLASQIVRELICVYMAKFVFAKTNDP--VESLCLFCFMTDRVDKTLGQDNFEE 1330
QY 1032 FLSSKRPVLVARDYKFLPCTVNLDPKL--EIIADVARARQOISIGLAVPPLPLHGG 1088
Db 1331 VARSK--DIEVLESGKPIYVDCYGNLADLTGGOOLVNFYSFKEN----- 1373
QY 1089 PRAPAGY-----SOPPSVCSSTSFNGPRAGVVSQPHSSYSGMTGQHPFPYNGSGP 1143
Db 1374 --RLPESIKIRDTSOEP--CGRLSF----- 1394
QY 1144 APGPVVLNLSINLVANDACEKLEGLDQSMLPQYTTIKKANINGVLAQCNIDELKEM 1203
Db 1395 -----LKEPKTKGLPQYAV-----CNINILPLHAKKAEKADRQ 1429
QY 1204 NMFQGMHLFRSTVLEMRNASHVPEDEPRFLSESSGAPAPGEPARR-----ASHN 1255
Db 1430 S-----FASIALKRRYS-----YILEPMSPO--SPCFERDIRMAIYADHL 1468
QY 1256 ELPHTELSSQTPYTLNPFEEELNTGLDGAHRHSLMS-----W--QOQRTPTBL 1304
Db 1469 GLSWTELARB--LNFVDEINQIRVEN--PNSLISOSFMLKKWTRDRKNNATDGL 1521
QY 1305 SLSNODSSISLKLTDKVOAEL-----RDAYREYIAOMSLGEGPG----- 1346
Db 1522 TVLTKINRDIYTLLEGPIFDYGNISGTRAFDENNVHDPVDRGHSQVELETFMGIX 1581
QY 1347 -----STTISGSSPHSTYYMGSSGGSIS--NLEQKKG-----KDS--BP 1385
Db 1582 WPPNPFOODDHFSDSISIESPFRPSRLSDGLVPOQNIIEHTGPPVYVTAEDTSLDS 1641
QY 1386 KPDDGKRSFLMKRGVYIDYSSSGVSTNDASPLDPTIEDEKSDQSSKL-----LPCK 1439

```

```

Db 1642 KMD-----SVTVTPADPLD--VDESOLKLCQSECAQCMASVPGIP 1682
QY 1440 SSERS--LFOUCLKGSGLYOKLPSPDESGTE-----ESDNFTLLKXDK 1485
Db 1683 NGROAEPLRPTKRVKGSSEQOEKKGSPDEVEDYKSLFEDYOLEEVAEMTEDQ 1742
QY 1486 DRKAEGKVERVP-----KSPENSAE--PIPTFKAREYLSDALDKDSSGVR 1533
Db 1743 GQAMLRVQRAELAMSLAGMONEPTPSGLESPPAQ-----RLTGLGLDRLDSDQAR 1797
QY 1534 SS-----ESSRPHSLHNEVADDSOLEKANLIELEDSSGKRG1-----PH 1574
Db 1798 DSITSVLGEPEKIKANGNHTA--EVIPEA--KAKPYPPSQNDIGKQSIKENLKPKTH 1852
QY 1575 SLSGLODPIARMSICSEDKKSPSCS--LLASSPEEMMPACQKAYNLNFTPTVTLNN 1632
Db 1853 GCGRTREPV--SPLTAYOKLEETSKVIEDAKPCVPYGMK--KMTFTTAD----- 1900
QY 1633 SAPANRANONFDEMEGIRETSQVILRPSSPNTTIONENLKSMTK 1679
Db 1901 -----GKARLNLQEEEG-----STRSEPK--QEGGYKVKTKK 1930

RESULT 11
149502
ankyrin - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 13-Aug-1999
C:Accession: I49502
R:White, R.A.; Birkenmeier, C.S.; Peters, L.L.; Barker, J.E.; Lux, S.E.
Mamm. Genome 3, 281-285, 1992
A>Title: Murine erythrocyte ankyrin cDNA: Highly conserved regions of the regulatory
A:Reference number: I49502; MUID:9245717; PMID:1386265
A:Accession: I49502
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1862 <RBS>
A:Cross-references: GB:M4756; MID:g191939; PID:AAA37336.1; PID:g191940
C:Genetics:
C:Gene: Ank-1
C:Superfamily: ankyrin, ankyrin repeat homology
C:Keywords: alternative splicing
F:40-72/Domain: ankyrin repeat homology <AN01>
F:73-105/Domain: ankyrin repeat homology <AN02>
F:106-138/Domain: ankyrin repeat homology <AN03>
F:139-167/Domain: ankyrin repeat homology <AN04>
F:168-200/Domain: ankyrin repeat homology <AN05>
F:201-233/Domain: ankyrin repeat homology <AN06>
F:234-266/Domain: ankyrin repeat homology <AN07>
F:267-299/Domain: ankyrin repeat homology <AN08>
F:300-332/Domain: ankyrin repeat homology <AN09>
F:333-365/Domain: ankyrin repeat homology <AN10>
F:366-398/Domain: ankyrin repeat homology <AN11>
F:399-431/Domain: ankyrin repeat homology <AN12>
F:432-464/Domain: ankyrin repeat homology <AN13>
F:465-497/Domain: ankyrin repeat homology <AN14>
F:498-530/Domain: ankyrin repeat homology <AN15>
F:531-563/Domain: ankyrin repeat homology <AN16>
F:564-596/Domain: ankyrin repeat homology <AN17>
F:597-629/Domain: ankyrin repeat homology <AN18>
F:630-662/Domain: ankyrin repeat homology <AN19>
F:663-695/Domain: ankyrin repeat homology <AN20>
F:696-728/Domain: ankyrin repeat homology <AN21>
F:729-761/Domain: ankyrin repeat homology <AN22>
F:762-794/Domain: ankyrin repeat homology <AN23>

Query Match 5.9%; Score 525; DB 2; Length 1862;
Best Local Similarity 20.1%; Pred. No. 4.4e-21;
Matches 396; Conservative 272; Mismatches 648; Indels 654; Gaps 77;

```

```

QY 21 LKALLEKKNDVERDEGCGTPLMTAAEGNLEIYKEILKNGANCNLELDWMTLLISASK 80
Db 90 VRELNVGANCANVAOSQKGFPLYMAGDENHLEVKFLENGANONVATEDEGFTPLAVALQ 149

```



```

Db 1103 LLMGDELDSPPEELAGKRIICRITTKDF-----POYFAVVS--RIKOSNOQICE 1150
Qy 951 GRLLRAN-----QISFMMDRLASVINLLEQMPYRTSLIILETEBIPQOMLTYER 1005
Db 1151 GGLISSTTVPLVCAQSPFEGALTKRIKRGLO-----AQVPDEIKYKILGNK 1196
Qy 1006 ISKNIPYTKDVEPLELIDGDIRNEFVLSRTPYLVADYKVFLLPCTVNDPLKRLIAD 1065
Db 1197 AT-----FSPITVTEPRRRKHKRITMTIP----- 1222
Qy 1066 VRAAREOISIGGLAYPPPLPHEGPPRAPSGYSDP--PSVSSSTSFNGPAGVVPQPHS 1123
Db 1223 -----PPSGEGVSNCKGDTTPNRLKLSIT-----GGSP----- 1254
Qy 1124 SYSGMTGPQHPFYNRSGSPAPGVYLLN-----SLNVD-----AVCEKIKQIEGLDQM 1173
Db 1255 AQWEDITGT-----PLTIKCCVSEFTTNVSAKRFMLADHQVLEVTGLATQL 1301
Qy 1174 -----LPOYCTTIKKANINGRYLAOCNI-----DELKK--EMNMNFGDHLFRSTVLEM 1220
Db 1302 YRELICVPYMAKFVFAKMDPVESSLRCFCMTDDKDKTLLEQENFE-----VA 1352
Qy 1221 RNASHVVPEDPRELSSSSGSPAPHGEPARRA-----SHNELPHT-----ELSSQTPYT 1269
Db 1353 RSKDIEVLEKRPITY-DCYGNLAPLTGGOQLVFNFSFKENRPLFSIKIRDTISOEPCGR 1411
Qy 1270 LNFEEELNTLGLDEGAPRHSNLSMOSQTRTPSLNSODSSIEISKLTQVQVEXRD 1329
Db 1412 LSPFAKERKTKGLPQTVACMLNLTILPAHKKETES-----DQDEIE--KIDRKQSHASL 1463
Qy 1330 AYRE--YIAQMSOLEGPGSTTISGRSSPHSTY-----YMGSSSGGSIHNSLE 1376
Db 1464 ALRRKRYSYLTPEGMIEKSTGAT--RSLP--TYSYKPFSTRPYQSWTAPITV----- 1513
Qy 1377 QEKKDEPDPDCKRKFLLMRGVIDYSSSGYSTNDASPLDPI-----TEDEKSDQSGS 1432
Db 1514 -----PGPAKSGF-----TSLSSSSMTSPASPLKSIWSVSTSPISKSTLGAS 1556
Qy 1433 KLIPGKKSSERSLFDLDLKLKSGSLR--YOKLPSEDESGT-----EESDMPPL----- 1480
Db 1557 TTSSVKISIDVAPSIIR-LRTMSSPIKTVYQSOPYNIQVSSGTLARPAVTEATPLKGLA 1615
Qy 1481 -----LKDDKRAKAEKVER-----VPSPE-----HSAEPITFTI-KAEKYLSDAL 1521
Db 1616 SNSTFSRTSPVTAGSLRERSITMTPPASPKSNIMYSSSLPFSIITSAPLSSPL 1675
Qy 1522 -----LDKSSDSGVSSESSESPNLSL--HNHVA-----DDSOLEKAN--LIE-- 1560
Db 1676 KSVVSPYKSRVDIYSSAKITMASSLSPPVKOMFGHAEVALVNGSISPLKASSSTLINC 1735
Qy 1561 -----LEDDSHSGKRGIPHSLSGLDPIIARMSICSEDKSPSECSLIASSPENMPAQ 1615
Db 1736 KATATLOEKISSATNSSVYSAATDVEVEFTTAMPSPILR-SYVSAP----- 1786
Qy 1616 KAVNLANFTBSTVTLNNSNAPANANONFDEMEGIRETSQVILRPSSS-----ENPTTTON 1670
Db 1787 SAFOSLRTPSASALYSLGSSISBATIS-----SVTSSIIITVPYSVVAVLEPAPLKL 1839
Qy 1671 ENLKSMTKRSQSS--YTRLKSDPPELHMAAS 1702
Db 1840 PDSNSFTKSAALLSPKILTLTETHPOPHSRTS 1873

```

RESULT 14 T42714

ankyrin 3, splice form 2 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000
C:Accession: T42714
R:Peterson, L.L.; Johnson, K.M.; Lu, F.M.; Elcheer, E.M.; Higgins, A.; Yialamas, M.; Turtzo, J.
J. Cell Biol. 130, 313-330, 1995
A:Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin gene
the repeat domain.

```

A:Reference number: Z22237, MUID:95340633, PMID:7615634
A:Accession: T42714
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1765 <PEP>
A:Cross-references: EMBL:L40632; NID:g710548; PID:g710551; PIDN:AAB01605.1
A:Experimental source: strain C57BL/6J; kidney
C:Genetics:
A:Gene: Ank3
A:Map position: 10
A:Introns: 1587/1
C:Superfamily: ankyrin; ankyrin repeat homology
C:Keywords: alternative splicing

Query Match      5.8%; Score 516; DB 2; Length 1765;
Best Local Similarity 19.8%; Pred. No. 1.3e-20;
Matches 408; Conservative 253; Mismatches 636; Indels 762; Gaps 83;

Qy 14 EENINIPALKLEKCDVDERNBCCGTPIMIAEONGNEIYKELIKNGANCULEDDMWT 73
Db 66 KEGHVEVVELLDQREANVDAATKKGNTALHIASLAGQAEVAVLVLTNGANVAASONGPT 125
Qy 74 ALISAKEGHVIHVEELKCGVNLERHDMGWTALMMACYKGRDVEELL----- 124
Db 126 PTYMAQENHLEVREFLLDNGASQSLATEDEGTPPLAVLQGHQDVVSLLENDTKKVR 185
Qy 125 -----SIGANPSYGL----- 135
Db 186 LPAHLIAARKDPTKAAALLLNDNTADYVESKSGFTPLHIAHYGINATLILNRAAAD 245
Qy 136 ---QVSVYPLIIMAGRGADIVHLLDNGCAKNCSDKGYTPLVMAARKG----- 182
Db 246 FPARNDITPLHASKRGNANMKLLIDRAKIDATROGTLPLHCGAASGHEQVEMLLD 305
Qy 183 -----HLECYHLLAMGADVDOEGANSMTALYAVKGYTQS 219
Db 306 RSAPILSKTKNGSLPLHMAQGDHLNCVQLLQHNVPVDVYNDYTLALHVAHCGHYKV 365
Qy 220 VREILLKRNPNVLTLDGNTALMIASKE----- 247
Db 366 AKVLDKRAKSPNAKALNGFTPLHIAKKNIRIMELLKNGASIQAVTESGILPIHVAAF 425
Qy 248 -GHTTEIVODLDAGYVNIIPDRSGDPLVLGAVRGHVEIVARLLQYADIDIRGDNKTA 306
Db 426 MGHVNIYVSLDLMHNGASPNNTNVRGELHMAARSGAEVYKLYVDGQAEKAKADDOTP 485
Qy 307 LTVAVEKGNATVRILOCPNPTETCTYKDGFTPLIKATKRNIEVEELLDKGAKVASVD 366
Db 486 LHSARLGRADIVQOQLDQGASPNNAATTSGYTPLHIAAREGHEVAAEFLDHGASLSITT 545
Qy 367 KKGDPRIHAIAGRSKLAELLRNPKGRLLYRNNKAGEPTYNIDCSHOKSILQITGA 426
Db 546 KKGFTPLHVAAYTGKLEVASLLOKSAS-----PDAAG-----KSGILPLIYA 588
Qy 427 RHLSPTETVDGMLGYDLYSALADILSEPTQPIPCVGLYAQWGSKSFLLKLDEDKT 486
Db 589 AH-----YDNCKVAL-----LLDDQGSPIRAAK----- 612
Qy 487 FNGQOIEPLPQPSWLVFLTLILLCGLLELFAFVHPNIGLAVSSTALLYEFTIYIF 546
Db 613 -----NGYTPLHIAAKNOMDIATSL-----LEY 636
Qy 547 GG-----REGESWMAWVLSRRLARHIGYLELLKLHFNVPPELPEQTKALPVRLEPT 601
Db 637 GADANAATVROG-----IASYHLAAQEDHVMVLLLSRN-----A 671
Qy 602 DYNRLSVGGETSLAEMITATISDACEREGFATLRFVFTYE-----DROGKKRW 652
Db 672 NVN-LSNNSGLRPL-----HLAQAEDVNVAEVLVNGAHVDAQTKRGY 714
Qy 653 KTKCCLPVIVFLFIIGCISSITLLAIFRVDPKHLITVNAVILISTASVVGALFVINCRTW 712
Db 715 TP-----LHVGHGYGNIKI-----VNFILQHSAKV-----NAKT- 743

```

```

QY 713 MOVLDSLNSQRRKLHNAASKLHKLKSGFMKVLKCEVLMARAKTITDSTQOTRLV 772
Db 744 -----KNGYTAALHQAQOQHITHINV-----LQNNASPMELTVNGNTALAI 785
QY 773 IIDGLDACEODKVLQMDTVRLVLSKGFIAIFASDPHIIITKAINQNLNSVLRSDNIGH 832
Db 766 -----ARRLGIVSVTLKAV--TEELMTTTTITERIKM--NPEETMNEVLDMSD----- 831
QY 833 DYMRNIYHLPVFLNLSRGLSNARK-----FLVTSATNGDVPSCDPTTGIEDADRVSQNS 886
Db 832 DEVRKASAPKLSDEYISDEGEDKCTWFKIPKQOEVLVXSED--AITGDTDXLGPQD 889
QY 887 LGEM-----TLGSKTA-LNRDQYRRQMQRTITRQMSFPLTKLV----- 927
Db 890 LKELDDSLPAEGYVGSLSARSLRFSDDSTYLLRSSYARDSMIEELVPSKEQH 949
QY 928 ---TEDWFSDISPQTRRLNIVSVTGRLLRANOISFNW--DRLASWINTLQMPYRTSW 982
Db 950 LFTREPDSD--SLRH-----YSWADTLDN-VNLVSS-PVHSGF 985
QY 963 LILYEETRGIPDQNTLTYERISKNIPTKQVPELLEIDGDIRNEVFLSKRPVIVA 1042
Db 966 LVSFVWDARG-----GSMR-----GSR-----H 1003
QY 1043 RDVFLPCTVNLDPKLEIILADVAREQISIGLAVPLPLEHGPAPRPSGYSQPSV 1102
Db 1004 HEMRIITP-----PKCTAPRPIITCRLVYKRHLAMP--PMWGBELASRLVEMGA- 1053
QY 1103 CSSTFNGPFPAGVVPQPHSSY-----SGMTGPOHPFYNGSGAPGEPV 1149
Db 1054 -----GAOFLGPVIVEIPHEGSMGRERELIVLSENGETWKEHQFSKEDLAE----- 1103
QY 1150 LLSNLVANDAVCKLQIISGLQSM-----POYCTIKKANINGVLAQCNDELK 1201
Db 1104 LINGMD-----EELDSPELGTRICRITTYDPOYFAVVS-----TKO 1143
QY 1202 EMMNFGDWHLFRSTVLEMRNA-----BSHWPEP--PRFL-SESSSGPA 1243
Db 1144 EBNQIGREGILSSTVPLVQASPEGALTRIRVGLAQVPEETVAKIIGNKATFPI 1203
QY 1244 PHGEPARASHELPH-----ELSS 1264
Db 1204 VTVEERRRRKFRPIMTIPVPPSGEGVNGYKGDATPMLRLCITGTSAPAWEDTIG 1263
QY 1265 QRP-----YTLNFS-----FEELMTIGDEBAPRHSNLSMOSQTR-----DPS 1303
Db 1264 TTPPLTFIKCVSFTTNVSAEFLADCHQVLETVGL-----ASQLYRELICVY 1311
QY 1304 LSSL-----NSQDSIEISKLTD-----KVOAERYDAYREYIAQMSOLEGGPGSTT 1349
Db 1312 MAKFWVPAKTNPVESSLRRCFCMTDQVDTKLEQENFEVARS--KQIEVLEGPVIVD 1369
QY 1350 ISGRSSPHSTYMGSSSG-----SIHSNLEQ-----KGDSEPRKDDGKSTPLMRG 1399
Db 1370 CYGNLAP-----LTKGGQOLVFNFYSFKENRLPFSIKIRDSQEP-CGRSLFLKE-- 1418
QY 1400 DVIDSSSGVSTNDASPLD-----PIEEDEKSD--OSGSKLLPGKK-----SEERSL 1446
Db 1419 ---PRTTKGIPOTANCNLTITPAKKAKEKADROSFSALRKRYSTLTPSMSPOSFC 1475
QY 1447 FQTDLKL-----KSGGLRYQKLPSDEDESGTE-----ESDNT-----PLUK-----DDK 1485
Db 1476 EKTDIRMAIVADHLGSLWTELARELNFVDEINQIRVENPNLSLSQSMLLKKWVTRGK 1535
QY 1486 DKRAG-----KVERVP-----KSPBH 1502
Db 1536 NATTALTSVTLKIRKIDIVLLLEGPIDYGNISGTRSFADENNVFHDVPDQWONETPSG 1595
QY 1503 SAE-PIRPIFAKEYISDLDKXSSDSGVSS-----ESSPHSLHNEVAD 1549
Db 1596 SLESPPQA-----RRLTGGLDRDLDSSDQARDSTITSYLTGEPKGTIANGNHTA--EVIP 1648

```

```

QY 1550 DSQLEKANLIELEDSDSGKRGIT-----PHSLSGLODPITARMSCSEDKSPSECS- 1601
Db 1649 EA--KAKPYFPSSNDIGKOSIKENLKRTHCGCTEPP-----SPLTAIYKSLSEFSK 1701
QY 1602 -LIASSPENWPAQACAYNINRTPSTVTLNNSAPANRANQFEMEGIRETSQVILRS 1660
Db 1702 LVIEDAPKPCVPVGMK--KMTRTTAD-----GKARLNQOELEG-----S 1738
QY 1661 SSPNPPTIQENENKSMTHK 1679
Db 1739 TRSEPK--QGEYKVKTKK 1755

RESULT 15
S37771
ankyrin, erythrocyte - mouse
C:Species: Mus musculus (house mouse)
C:Date: 27-May-1994 #sequence_revision 03-Aug-1995 #text_change 13-Aug-1999
C:Accession: S37771
R:Birkemeier, C.S.; White, R.A.; Peters, L.L.; Hall, E.J.; Lux, S.E.; Barker, J.E.
J. Biol. Chem. 268, 9533-9540, 1993
A:Title: Complex patterns of sequence variation and multiple 5' and 3' ends are found
A:Reference number: S37771; MIMD:93252825; PMID:8486643
A:Status: preliminary
A:Accession: S37771
A:Molecule type: mRNA
A:Residues: 1-1848 <BIR>
A:Cross-references: EMBL:X69063; MIMD:g311816; PIDD:CAA48801.1; PID:g311817
C:Superfamily: ankyrin; ankyrin repeat homology
C:Keywords: alternative splicing
F:48-80/Domain: ankyrin repeat homology <AN01>
F:81-113/Domain: ankyrin repeat homology <AN02>
F:114-146/Domain: ankyrin repeat homology <AN03>
F:147-175/Domain: ankyrin repeat homology <AN04>
F:176-208/Domain: ankyrin repeat homology <AN05>
F:209-241/Domain: ankyrin repeat homology <AN06>
F:242-274/Domain: ankyrin repeat homology <AN07>
F:275-307/Domain: ankyrin repeat homology <AN08>
F:308-340/Domain: ankyrin repeat homology <AN09>
F:341-373/Domain: ankyrin repeat homology <AN10>
F:374-406/Domain: ankyrin repeat homology <AN11>
F:407-439/Domain: ankyrin repeat homology <AN12>
F:440-472/Domain: ankyrin repeat homology <AN13>
F:473-505/Domain: ankyrin repeat homology <AN14>
F:506-538/Domain: ankyrin repeat homology <AN15>
F:539-571/Domain: ankyrin repeat homology <AN16>
F:572-604/Domain: ankyrin repeat homology <AN17>
F:605-637/Domain: ankyrin repeat homology <AN18>
F:638-670/Domain: ankyrin repeat homology <AN19>
F:671-703/Domain: ankyrin repeat homology <AN20>
F:704-736/Domain: ankyrin repeat homology <AN21>
F:737-769/Domain: ankyrin repeat homology <AN22>
F:770-802/Domain: ankyrin repeat homology <AN23>

Query Match 5.78; Score 506; DB 2; Length 1848;
Best local similarity 19.44; Pred. No. 5.2e-20;
Matches 378; Conservative 284; Mismatches 653; Indels 634; Gaps 73;

QY 21 LKALLECKVDRENEGQGPLMIAEOGLLEVLKELIKGANCNLEDIDNTATLSAK 80
Db 98 VRELVTNGAVNNAQSQGFTPLMAOENHLEVKFLLENGANQNAVATEDGFTPLVALQ 157
QY 81 EGHVHVEELKGV-----NLEHRDMGWTALMA 111
Db 158 QGHENVAHLINLTGKGVRLPALHTAARNDDRTAAVLQNDPNPDVLSGTGFTPLAHT 217
QY 112 CYGRIDVVELLSHGANSVT-----GLQY-----SV 139
Db 218 AHENLVNAQLLNRGASVNFTPONGITPLHISRGVNIWVLLDRGAQIETRTKDEL 277
QY 140 YPIITMAGRHADIVHLLQNGAKVNCSDKYGTTPLVMAARGHLECYVHLLAMGADV 198
Db 278 TPLHCARNGNHVRISLTDHGAPIOAKTKNGLSPIHMAAQGDHCDVRLQYNAEIDD 337

```

```

QY 199 -----ORGA-----NSMTALIVAGGYTQSVKELTKR 226
Db 338 ITDLHLPRLVAAHCGHHRVAKVLLDKGAPNSRBNLNGTPLHACKNNIRWELLTKT 397
QY 227 NPNVNLTDKGNATLMAISKEGHEIVODLLDAGTYVNIIPRSDVTLIGAVGHEIY 286
Db 398 GASIDAVTESGLFPLHVASFMGHLPIYKNNLORGASPNVSNVVERPLHMAAAGHTEVA 457
QY 287 RALLQKADIDIRQDNKTALYMAVEGNATVYRDIQCPDTEICTKDETPPLIKATKM 346
Db 458 KYLLQNKAKANAARAKDQPLHCAARIGHGMVKLLENGASPNLATTAHTPLHAAARE 517
QY 347 RNEVEVELLDKGAASVADKGTPLHAIIRGSRKLAELLRNPDGSLLRPNKAGE 406
Db 518 GHVDTALALLEKESQACMTKKGFTPLHVAKATGKVALLELHDA-----HPNACK 571
QY 407 ---TPYNIDCSHO-----KSL-----TOIFGAR-----427
Db 572 NGLTPLHVAHNNHNDIVKLLRPGSGPSHPANNGYTPHLIAKQNOIEVARSLQYGS 631
QY 428 -----HSPETGDMHGYDLYSSALADIISEPTWOP-----PIC---462
Db 632 ANAESVGVYTPHLAAGEHTEMVALLSKQANGNKGSLTPLHLVSOEGHVPADVL 691
QY 463 -----VGLYAQWGSCK---SFLKKLED---EMKTFAGQJIEPLFOFSWL- 501
Db 692 IKHGVTYDATTRMGYTPHLVASHYGNKLYKFLQHOADVNAKTKG--YSPHQAQOQ 749
QY 502 -IYFLLTLLCGGLGLFAFTVHPN-----LGIAVSLSPALLYIFPIYIFGGRRE 551
Db 750 HTDITVLLKNG-----ASPNEVSNCTPLALAKRGIYSIVDYLVKVI-----794
QY 552 GESNNAMVYSTRARHIGYLLLELLKLMFNPPELQOTTKALPY--RFLPFDPYPLSS 608
Db 795 -----DETSVVLVSDKHMSYPTETVDELIDVS---EBEGTAHISIMGDELIVGSKARRRS 846
QY 609 --VGETSLAEMIATISDACEREGFLATRLFR-----VEKTED--TQKKKKKKKTC 656
Db 847 RDVGEKEKELDPYKLDQVYE-----SPAIPRIPCVTPPTVYIRSDBOQASKEYDEDS 900
QY 657 CLPSEVLEFLFICIIISGITLLAIFRVDPKHLTVNANVLISIASVGLAFVLCRTWQVYL 716
Db 901 LIPS-----SPATERSDNI--SPVASPVHTGELVSP-----WV 931
QY 717 DSIILNSORRKLHNAASKLHKLKSEGFVKYKCEVELMARRAKTIIDSTONQPLVVIIDG 776
Db 932 DARGGSMRSGRHN-----GLRVYIP-----PRCAAPTRITCRLV-----966
QY 777 LDACEODKVLQMLDTPVAVLEFSKGPFLAIFASDPHIIIIKAINONLSVLRDSNINHDMR 836
Db 967 -----KQOKINTPPPL-----AEEGLASRIIILGPTGAOFUS 999
QY 837 N-IVHLPEVL-NSRGLSNARKFLVTSATNGVPCSDTGTQEDADRRVSONSLGEMTKLG 894
Db 1000 PVIIEIPHFASHGRG---DRELVVLRSENSV-----WKHKRSRYGGSYLDQILN-G 1047
QY 895 SKTALNRDQTYRRRQOMORTTTRQMSFDLTKLVY---EDWFSDISPOJMRRLINIVSVTG 951
Db 1048 MDEELISLELEKERRCRIT--TDPELVFVIMSRLOC-DYDTIGPEG-----GSLRS 1097
QY 952 RLLEANOISFNMRCLASMINLITQOWPYRTSMWLLYLEETEGIPDOMTLTIYERISKNTIP 1011
Db 1098 KIVPLVQATPEPNAVINKKVALQ-----AOPVDELVTYTKL-----1134
QY 1012 TTADVPELLEIDGIRNEFEVFLSSRTPVLVARDVKVFLPCTVNLDEKREIITADVPARE 1071
Db 1135 -----GNQATE-----SPI-----VYVEPRRRKFRHPI-----1157
QY 1072 QISIGGLAVPPLPLHSGPPRAPBSGYISOPS---VCSSTSPNGFPAGVAVSPDHSYSYG 1128
Db 1158 -----GLRIPLPSPWTDNBR-DSGEGDITSLRLCS-----VIGTDQAOQ-----WED 1199

```

```

QY 1129 MTGQHPHFYNR-----GSGPAPGVYVNLNSL---NVDANCEKL-----Q 1165
Db 1200 ITGTYKLITIANECANFTTVNSARFMLSQCPRTAEVHRATLTKELTAVPYMAKVIYPAK 1259
QY 1166 IEGLOSMPLPOYCTTIKK-----ANINGRVLAQONI 1196
Db 1260 MNDARGRLRCYCMTDKDKYDKTLEQHENVEVARSDIEVELEGMPLFAELSG-----NL 1313
QY 1197 DELKEMANNFGDWHLEFRSTVLEMRNASHVVEDEPRELSESSSGPAPGEPARASHNE 1256
Db 1314 VPYKKAQOORSEHFQSFRENRLAI-----PVKYRDSRREPGEPLSPFRKTKYE 1362
QY 1257 -----LPHTELS-----SOTPTIINFPELNTLGLDEGAPR-----1288
Db 1363 DTQHILCHLNTIMPCTKSGAEDRRRTTLPLTLRYSLISESLGFTSDTDVEMRMAYI 1422
QY 1289 ---HSNLSWOSQTR-----RTPSLSLNSQDSI-----EISKLTDKVOA 1325
Db 1423 REHLGSLMAELARELOFSVEDINRIKVENPNSLDQSTALLTLWYDREGENAKM-ENLYT 1481
QY 1326 EYRDAYREYIAQMSOLEGGPGSTTISGRSSPHSTYVMQOSSGSGSIHSLNLEQEKDSEP 1385
Db 1482 ALRNIIDRSEIVMLE-----VSGROS-----RNLKP 1507
QY 1386 KPDDGRKSLMKRGVYIDYSS-----SGVSTNDASPLDPTIEDEKSDQSGSKILLPKK 1439
Db 1508 ERHGGREYSLSPQVNGYSSLODELLSPASLOVALP-SPLCADOYMEVYITDAIPLAA 1566
QY 1440 SBRSSLFQTDLKLKSGGLRYOKLPSEDE---SGTEESDNPPLKDKDKRAEKGVERV 1496
Db 1567 TEHDPTLEKSDMQVMSAGITPLSLVYAEBSLSECSKAEDSDALP-----EMKLEGAHSED 1620
QY 1497 PKSPHSAEPT---RTFIKAEVYSDALLDK-----DSSDSGVRSSES 1537
Db 1621 TQPELGSODLVEDYTVSDATNGLADLIGQORVHARITDPSVQVLDRSQARTLDWDK 1680
QY 1538 SPNHSIHNEVADDSOLEKANLIELEDHSHSGRGI-----PHLSLGLQDPIIARMSIC 1590
Db 1681 QOSTAVHPEAOSSMOE---EVQGPHSFQGRITTTIQQPRPALOEYEDLV---STR 1733
QY 1591 SEDKKSPECSLJASSPEENMPACOKAYN 1619
Db 1734 EHVQGRPPPTGSPKAGKEPSLMAPESSAP 1762

```

Search completed: July 1, 2003, 14:43:12
 Job time : 79 secs

